

A Trellis Coded Modulation Scheme for the Fading Relay Channel

Vijayvaradharaj T Muralidharan
Dept. of ECE, Indian Institute of Science
Bangalore 560012, India
Email: tmvijay@ece.iisc.ernet.in

B. Sundar Rajan
Dept. of ECE, Indian Institute of Science,
Bangalore 560012, India
Email: bsrajan@ece.iisc.ernet.in

Abstract—A decode and forward protocol based Trellis Coded Modulation (TCM) scheme for the half-duplex relay channel, in a Rayleigh fading environment, is presented. The proposed scheme can achieve any spectral efficiency greater than or equal to one bit per channel use (bpcu). A near-ML decoder for the suggested TCM scheme is proposed. It is shown that the high SNR performance of this near-ML decoder approaches the performance of the optimal ML decoder. The high SNR performance of this near-ML decoder is independent of the strength of the Source-Relay link and approaches the performance of the optimal ML decoder with an ideal Source-Relay link. Based on the derived Pair-wise Error Probability (PEP) bounds, design criteria to maximize the diversity and coding gains are obtained. Simulation results show a large gain in SNR for the proposed TCM scheme over uncoded communication as well as the direct transmission without the relay. Also, it is shown that even for the uncoded transmission scheme, the choice of the labelling scheme (mapping from bits to complex symbols) used at the source and the relay significantly impacts the BER vs SNR performance. We provide a good labelling scheme for 2^l -PSK signal set, where $l \geq 2$ is an integer.

I. PRELIMINARIES AND BACKGROUND

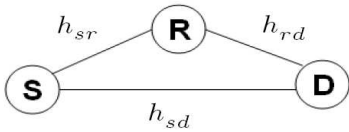


Fig. 1. The Relay Channel

We consider the Rayleigh fading relay channel shown in Fig. 1, consisting of the source node S , the relay node R and the destination node D . It is assumed that R can operate only in the half-duplex mode, i.e., it cannot receive and transmit simultaneously in the same frequency band. It is assumed that R has perfect knowledge about the instantaneous value of the fade coefficient associated with the S-R link and D has perfect knowledge about the instantaneous values of the fade coefficients associated with the S-R, R-D and S-D links. Throughout, the phase during which the relay is in reception mode is referred to as Phase 1 and the phase during which the relay is in transmission mode is referred to as Phase 2.

In all practical scenarios the S-R and R-D links are stronger than the S-D link. As a result, the use of coding schemes which involve both S and R , can potentially outperform the

coding schemes involving S alone. The problem addressed in this paper is the design of a Trellis Coded Modulation (TCM) scheme for the relay channel, which achieves a spectral efficiency greater than or equal to one bit per channel use (bpcu) and provides a large gain compared to the TCM scheme for the direct transmission without the relay.

A comparison of the proposed TCM scheme is made with the uncoded transmission scheme for the relay channel operating in the Decode and Forward (DF) mode, in which no coding is done at S and R , which is described as follows: At S , uncoded bits are directly mapped onto complex symbols from a signal set during Phase 1 and transmitted. During Phase 2, S and R map the uncoded bits (S uses the same bits as in Phase 1 and R uses the bits it decoded during Phase 1) onto complex symbols from a signal set for transmission. The signal sets used at S during Phase 1, S during Phase 2 and R during Phase 2 are assumed to be the same, but the mapping from bits to complex symbols can be different.

A. The Proposed Scheme

The proposed TCM scheme for the fading relay channel is shown in Fig. 2(a) and Fig. 2(b) (shown at the top of the next page). It is assumed that TCM encoding at S during Phase 1 and Phase 2 and at R during Phase 2 take place using the same encoder, but the labelling schemes used for mapping coded bits onto signal points can be different. By a trellis, we refer to the states of the TCM encoder, the edges connecting these states in two successive stages and the complex numbers which are labelled on these edges. By a path in the trellis, we refer to the sequence of connected edges (the complex numbers labelled on the edges are not included). Let \mathcal{T}_{S_1} , \mathcal{T}_{S_2} and \mathcal{T}_R represent the trellises used for encoding by S during Phase 1, by S during Phase 2 and by R during Phase 2 respectively. Since the encoder corresponding to the trellises \mathcal{T}_{S_1} , \mathcal{T}_{S_2} and \mathcal{T}_R are the same, they differ only in the complex numbers labelled on the edges. Let ζ denote the set of edges connecting the states in two successive stages of the trellis. The labelling of the edges in the trellis \mathcal{T}_{S_1} is given by the map $\mathcal{X}_{s_1} : \zeta \rightarrow \mathcal{S}$, where \mathcal{S} is the signal set used at S and R . Similarly, the maps \mathcal{X}_{s_2} and \mathcal{X}_r are defined for the trellises \mathcal{T}_{S_2} and \mathcal{T}_R respectively.

During Phase 1, input bits at S are encoded by the TCM encoder, mapped onto complex symbols from the signal set \mathcal{S} and interleaved by the block interleaver as shown in Fig.

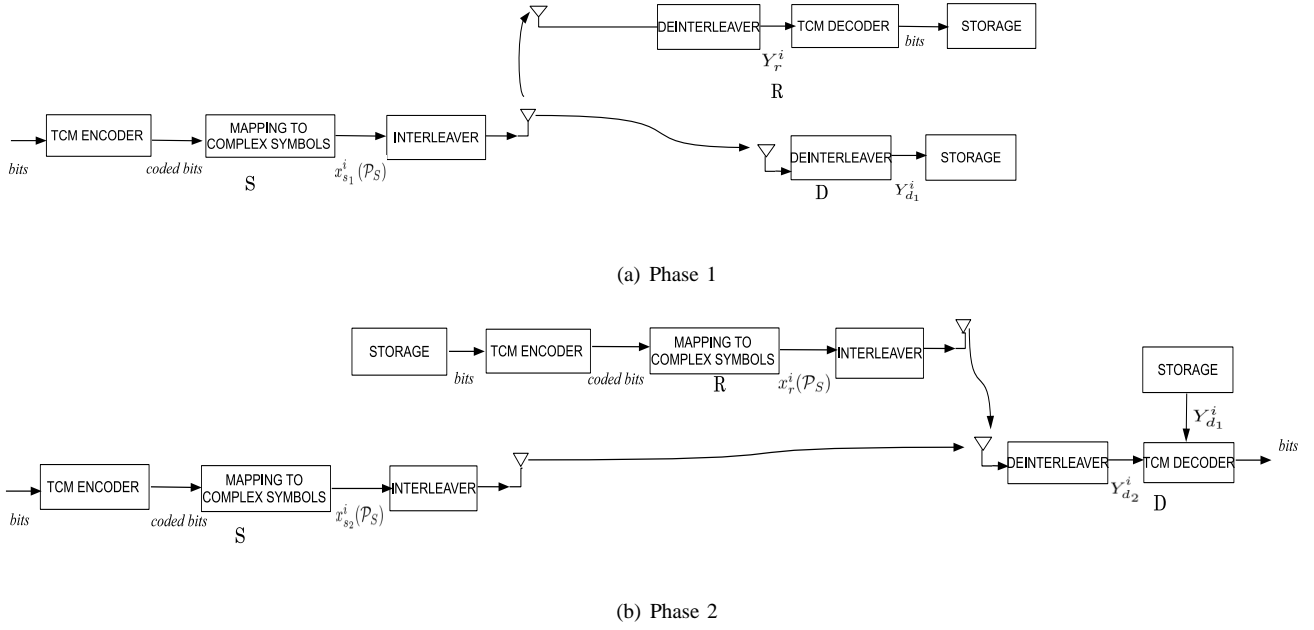


Fig. 2. The TCM Scheme

2(a). Each transmission made is assumed to be a block of L complex symbols. After each transmission, the encoder state is brought back to the all zeros state by appending zero input bits at the end. Let \mathcal{P}_S denote the path in the trellis corresponding to the output of the TCM encoder at S. Let $x_{s1}^i(\mathcal{P}_S)$ denote the symbol to be transmitted corresponding to the i^{th} branch of the path \mathcal{P}_S , where $1 \leq i \leq L$.

The received complex signal sequence at R and D during Phase 1 are given by,

$$\left. \begin{aligned} Y_r^i &= h_{sr}^i x_{s1}^i(\mathcal{P}_S) + z_r; \\ Y_{d1}^i &= h_{sd1}^i x_{s1}^i(\mathcal{P}_S) + z_{d1}; \end{aligned} \right\} 1 \leq i \leq L$$

where h_{sr}^i and h_{sd1}^i are the independent zero mean circularly symmetric complex Gaussian fading coefficients associated with the S-R and S-D links respectively with the corresponding variances given by σ_{sr}^2 and σ_{sd}^2 . The additive noises z_r and z_{d1} are assumed to be $CN(0,1)$, where $CN(0,1)$ denotes the standard circularly symmetric complex Gaussian random variable. An interleaver with sufficiently large block length is assumed to make the block fading channel appear as a fading channel with independent fade coefficients for successive channel uses. At R, after deinterleaving, an ML decoder using the Viterbi algorithm is used to decode the bits (Fig. 2(a)). Let \mathcal{P}_R denote the path in the trellis to which R decodes. At D, the received complex numbers are deinterleaved and stored (Fig. 2(a)).

During Phase 2, S encodes the same bits used during Phase 1 and R encodes the bits it decoded during Phase 1, using the TCM encoder. The encoded bits at S and R are mapped onto complex symbols from the signal set \mathcal{S} (possibly with different mappings and also possibly different from the mapping used at S during Phase 1) and interleaved using the

block interleavers (same as the interleaver used at S during Phase 1) as shown in Fig. 2(b). Let $x_{s2}^i(\mathcal{P}_S)$ and $x_r^i(\mathcal{P}_R)$ denote the symbols transmitted by S and R respectively during Phase 2, corresponding to the i^{th} branch of the paths \mathcal{P}_S and \mathcal{P}_R , where $1 \leq i \leq L$.

The received complex signal sequence at D during Phase 2 is given by,

$$Y_{d2}^i = h_{sd2}^i x_{s2}^i(\mathcal{P}_S) + h_{rd}^i x_r^i(\mathcal{P}_R) + z_{d2},$$

where $1 \leq i \leq L$. The random variables h_{sd2}^i and h_{rd}^i are the independent zero mean circularly symmetric complex Gaussian fading coefficients associated with the S-D and R-D links respectively with the corresponding variances given by σ_{sd}^2 and σ_{rd}^2 . The additive noise z_{d2} is $CN(0,1)$. Throughout, it is assumed that $\sigma_{sd}^2 \ll \sigma_{rd}^2$, due to the proximity of the relay to the destination than the source to the destination.

During Phase 2, at D the received complex numbers are deinterleaved and along with the complex numbers stored during Phase 1 are fed to the proposed near-ML TCM decoder described in Section III.

B. Background and Related Work

Achievable rates and upper bounds on the capacity for the discrete memoryless relay channel for the full-duplex and the more practical half-duplex relay channel were obtained in [1], [2]. These results were extended for the half duplex Gaussian relay channel in [3]. Optimal power allocation strategies to maximize the achievable rate for the Rayleigh fading relay channel were investigated in [4]. Several protocols like Amplify and Forward (AF), Decode and Forward (DF), Compress and Forward (CF) were proposed [5]. Power allocation strategies for the Non Orthogonal DF (NODF) protocol

were discussed in [6]. Non-orthogonal relay protocols offer higher spectral efficiency when compared with orthogonal relay protocols [7], [8], [9]. Hence, the proposed TCM scheme in this paper is based on the NODF protocol.

A turbo-coding scheme for the full duplex and more practical half-duplex relay channels were proposed in [10], [11]. These coding schemes achieve a spectral efficiency strictly less than one bpcu, i.e., they pay in bandwidth. Coding schemes for the relay channel with bandwidth constraints, which achieve a spectral efficiency greater than or equal to one bpcu are not reported in the literature, to the best of our knowledge. In this paper, we propose a Trellis Coded Modulation (TCM) scheme for the half-duplex relay channel for any spectral efficiency greater than or equal to one bpcu.

Implementation as well as the performance analysis of the optimal ML decoder for the relay channel operating in the DF mode is complicated [12], [13]. To solve these problems, several sub-optimal decoders for uncoded DF scheme were proposed in [12], [13], [14] and [15]. These problems carry over to coded communication using DF protocol as well. If the decoder at D is designed assuming that R always decodes the bits correctly, the diversity order obtained is dependent on whether the relay decoded correctly or not. To circumvent this problem, relay selection strategies [16], fountain codes [17], Cyclic Redundancy Check (CRC) codes [18] etc. are used. To solve these problems of finding a decoder with implementable complexity and tractable performance analysis, we propose a near-ML decoder, whose performance at high SNR approaches the performance of the optimal ML decoder. The form of this decoder allows a DF scheme in which the relay need not check whether it has decoded correctly before forwarding the message. This eliminates the need for embedding CRC bits as well as the need for feedback from the relay to the source. Furthermore, the diversity order of the proposed scheme does not depend on whether the relay decoded correctly or not.

The proposed TCM scheme is different from the Cooperative Multiple Trellis Coded Modulation (CMTCM) scheme proposed in [18] in the following aspects:

- The presence of embedded CRC bits and feedback from R to S is assumed in [18], whereas we make no such assumptions.
- Interleaving/Deinterleaving is not assumed in [18] whereas we assume interleaving/deinterleaving which makes the quasi-static fading scenario appear as a fast fading scenario.

C. Contributions and Organization

The main contributions of this paper are as follows:

- A TCM scheme for the fading relay channel is proposed which achieves any spectral efficiency greater than or equal to one bpcu.
- A near-ML decoder is obtained for the proposed TCM scheme. The relay need not check whether it has decoded correctly before forwarding the message which eliminates the need for embedded CRC bits as well as the feedback from the relay to the source.

- The proposed near-ML decoder enables the formulation of design criteria to maximize the diversity order and the coding gain.
- Bounds on the PEP are derived for the near-ML decoder based on which the criteria to maximize the diversity order and coding gain are obtained.
- The BER vs SNR performance of the near-ML decoder proposed with a non-ideal S-R link, at high SNR, approaches the performance of the optimal ML decoder with an ideal S-R link. This implies that the high SNR performance of the near-ML decoder for the relay channel with a non-ideal S-R link approaches the performance of the optimal ML decoder.
- It is shown that even for the uncoded transmission scheme, a proper choice of the labelling at S and R provides a significant performance improvement. We give a good labelling scheme for 2^l -PSK signal set, where $l \geq 2$ is an integer.
- Simulation results show a large gain in the BER vs SNR performance for the proposed TCM scheme over the uncoded transmission scheme as well as the TCM scheme for the direct transmission without the relay.

The organization of the rest of the paper is as follows: In Section II, the bounds on the ergodic capacity of the fading relay channel with Gaussian input alphabet are compared with the capacity bounds when the input is constrained to take values from a finite signal set. The description of the near-ML decoder for the TCM scheme constitutes Section III. In Section IV, the PEP expressions are derived, based on which the code design criteria to maximize the diversity order and coding gain are obtained. In Section V, effect of the choice of the labelling scheme for the uncoded transmission scheme is discussed and a good labelling scheme for 2^l -PSK constellation is presented. In Section VI, TCM code design examples are presented. Simulation results are presented in Section VII.

Notations: For a random variable X_s which takes value from the set \mathcal{S} , we use $x_{s,i}$ to represent the i -th element of \mathcal{S} . $\mathbb{E}_z[Y]$ denotes the expectation of Y with respect to the random variable z . Throughout, \log refers to \log_2 and $C(a)$ denotes $\log(1 + a)$. Let $CN(0, I_n)$ denote the standard circularly symmetric complex Gaussian random vector of length n . Also let $N(0, c)$ denote the scalar real valued Gaussian random variable with mean zero and variance c . For simplicity, distinction is not made between the random variable and a particular realization of the random variable, in expressions involving probabilities of random variables. In some probability expressions involving conditioning of the fading coefficients, the fact that the probability is conditioned on the values taken by the fading coefficients is not explicitly written, as it can be understood from the context. For a set \mathcal{A} , $|\mathcal{A}|$ denotes the cardinality of \mathcal{A} . For two sets \mathcal{A} and \mathcal{B} , $\mathcal{A} \cup \mathcal{B}$ denotes the union of the sets \mathcal{A} and \mathcal{B} . The superscript $(\cdot)^T$ denotes the transpose operation. The set $\{x \mid c(x)\}$ denotes the set of all values of x for which the condition $c(x)$ is satisfied. $\Re(x)$ denotes the real part of the complex number

$$R_1 \triangleq I(X_{s_1}; Y_r) = \log(M_{s_1}) - \frac{1}{M_{s_1}} \sum_{i_1=0}^{M_{s_1}-1} \mathbf{E}_{z_r} \left[\log \left(\frac{\sum_{i=0}^{M_{s_1}-1} \exp(-|z_r - h_{sr}x_{s_1,i} + h_{sr}x_{s_1,i_1}|^2)}{\exp(-|z_r|^2)} \right) \right] \quad (1)$$

$$R_2 \triangleq I(X_{s_2}; Y_{d_2} | X_r) = \log(M_{s_2}) - \frac{1}{M_{s_2}} \sum_{i_1=0}^{M_{s_2}-1} \mathbf{E}_{z_{d_2}} \left[\log \left(\frac{\sum_{i=0}^{M_{s_2}-1} \exp(-|z_{d_2} - h_{sd}x_{s_2,i} + h_{sd}x_{s_2,i_1}|^2)}{\exp(-|z_{d_2}|^2)} \right) \right] \quad (2)$$

$$R_3 \triangleq I(X_{s_1}; Y_{d_1}) = \log(M_{s_1}) - \frac{1}{M_{s_1}} \sum_{i_1=0}^{M_{s_1}-1} \mathbf{E}_{z_{d_1}} \left[\log \left(\frac{\sum_{i=0}^{M_{s_1}-1} \exp(-|z_{d_1} - h_{sd}x_{s_1,i} + h_{sd}x_{s_1,i_1}|^2)}{\exp(-|z_{d_1}|^2)} \right) \right] \quad (3)$$

$$R_4 \triangleq I(X_{s_2}, X_r; Y_{d_2}) = \log(M_{s_2} M_r) - \frac{1}{M_{s_2} M_r} \sum_{i_1=0}^{M_{s_2}-1} \sum_{j_1=0}^{M_r-1} \mathbf{E}_{z_{d_2}} \left[\log \left(\frac{\sum_{i=0}^{M_{s_2}-1} \sum_{j=0}^{M_r-1} \exp(-|z_{d_2} - h_{sd}x_{s_2,i} - h_{rd}x_{r,j} + h_{sd}x_{s_2,i_1} + h_{rd}x_{r,j_1}|^2)}{\exp(-|z_{d_2}|^2)} \right) \right] \quad (4)$$

$$R_5 \triangleq I(X_{s_1}; Y_r, Y_{d_1}) = \log(M_{s_1}) - \frac{1}{M_{s_1}} \sum_{i_1=0}^{M_{s_1}-1} \mathbf{E}_{z_r, z_{d_1}} \left[\log \left(\frac{\sum_{i=0}^{M_{s_1}-1} \exp(-|z_r - h_{sr}x_{s_1,i} + h_{sr}x_{s_1,i_1}|^2 - |z_{d_1} - h_{sd}x_{s_1,i} + h_{sd}x_{s_1,i_1}|^2)}{\exp(-|z_r|^2 - |z_{d_1}|^2)} \right) \right] \quad (5)$$

x . Throughout, E_S denotes the average energy in dB of the signal set \mathcal{S} used at the source and the relay. Let $Q[\cdot]$ denote the complementary CDF of the standard Gaussian random variable.

II. INFORMATION THEORETIC LIMITS

Throughout this section, the achievable rate of the decode and forward scheme is taken to be the lower bound on the capacity and the cut-set bound [1] is taken to be the upper bound on the capacity.

For the half duplex relay channel, the received signal at R during Phase 1 is given by,

$$Y_r = h_{sr}X_{s_1} + z_r.$$

The received signal at D during Phase 1 and phase 2 are given by,

$$\begin{aligned} Y_{d_1} &= h_{sd}X_{s_1} + z_{d_1}; \\ Y_{d_2} &= h_{sd}X_{s_2} + h_{rd}X_r + z_{d_2}. \end{aligned}$$

where X_{s_1} and X_{s_2} denote the symbols transmitted by S during Phase 1 and Phase 2 respectively, X_r denotes the symbol transmitted by R during Phase 2. The random variables z_{d_1} , z_{d_2} and z_r are independent and $CN(0, 1)$.

For the relay channel with Gaussian input alphabet, the lower and upper bounds on the capacity are given by [3] [4],

$$C_L^G = \frac{1}{2} \max_{0 \leq \beta \leq 1} \min \left\{ \mathbf{E}_{h_{sr}} [C(|h_{sr}|^2 E_S)] + \mathbf{E}_{h_{sd}} [C((1-\beta)|h_{sd}|^2 E_S)], \right. \\ \left. \mathbf{E}_{h_{sd}} [C(|h_{sd}|^2 E_S)] + \mathbf{E}_{h_{sd}, h_{rd}} [C(|h_{sd}|^2 E_S + |h_{rd}|^2 E_S + 2\sqrt{\beta}|h_{sd}||h_{rd}|E_S)] \right\}; \quad (6)$$

$$C_U^G = \frac{1}{2} \max_{0 \leq \beta \leq 1} \min \left\{ \mathbf{E}_{h_{sd}, h_{sr}} [C((|h_{sd}|^2 + |h_{sr}|^2) E_S)] \right. \\ \left. + \mathbf{E}_{h_{sd}} [C((1-\beta)|h_{sd}|^2 E_S)], \right. \\ \left. \mathbf{E}_{h_{sd}, h_{rd}} [C(|h_{sd}|^2 E_S + |h_{rd}|^2 E_S + 2\sqrt{\beta}|h_{sd}||h_{rd}|E_S)] \right. \\ \left. + \mathbf{E}_{h_{sd}} [C(|h_{sd}|^2 E_S)] \right\}. \quad (7)$$

To compute the achievable rate bounds with finite input constellation, assume $X_{s_1} \in \mathcal{S}_{s_1}$, $X_{s_2} \in \mathcal{S}_{s_2}$ and $X_r \in \mathcal{S}_r$, where \mathcal{S}_{s_1} , \mathcal{S}_{s_2} and \mathcal{S}_r are signal sets such that $|\mathcal{S}_{s_1}| = M_{s_1}$, $|\mathcal{S}_{s_2}| = M_{s_2}$ and $|\mathcal{S}_r| = M_r$. Define $R_4 = I(X_{s_1}; Y_r)$, $R_5 = I(X_{s_2}; Y_{d_2} | X_r)$, $R_6 = I(X_{s_1}; Y_{d_1})$, $R_7 = I(X_{s_2}, X_r; Y_{d_2})$ and $R_8 = I(X_{s_1}; Y_r, Y_{d_1})$. It is assumed that coding schemes in which all the possible choices for (X_{s_1}, X_{s_2}, X_r) occur with equal probability only are of interest. This forces the marginal distributions of X_{s_1} , X_{s_2} , X_r to be independent and uniformly distributed.

The achievable rate bounds of the half duplex relay channel with finite input constellation can be shown to be [19],

$$C_L^{CC} = \min \left\{ \frac{1}{2} \mathbf{E}_{h_{sr}} [R_1] + \frac{1}{2} \mathbf{E}_{h_{sd}} [R_2], \right. \\ \left. \frac{1}{2} \mathbf{E}_{h_{sd}} [R_3] + \frac{1}{2} \mathbf{E}_{h_{sd}, h_{rd}} [R_4] \right\}; \quad (8)$$

$$C_U^{CC} = \min \left\{ \frac{1}{2} \mathbf{E}_{h_{sr}, h_{sd}} [R_5] + \frac{1}{2} \mathbf{E}_{h_{sd}} [R_2], \right. \\ \left. \frac{1}{2} \mathbf{E}_{h_{sd}} [R_3] + \frac{1}{2} \mathbf{E}_{h_{sd}, h_{rd}} [R_4] \right\}. \quad (9)$$

where R_1 - R_5 are given by (1)-(5), shown at the top of this page.

The expressions for the capacity bounds (6), (7), (8) and (9) can be evaluated using Monte-Carlo simulations. The plots for the E_S Vs Capacity bounds for the relay channel with Gaussian alphabet as well as 4-PSK and 8-PSK constellations, for $\sigma_{sd}^2=0$ dB, $\sigma_{sr}^2=15$ dB, $\sigma_{rd}^2=15$ dB, are shown in Fig. 3. Fig. 3 also shows the capacity with Gaussian alphabet and the constellation constrained capacity with 4-PSK signal set for the direct transmission (S-D link). From Fig. 3, it is clear that the use of relay can provide significant advantage over the direct transmission. From Fig. 3, we can see that for a spectral efficiency of 1 bpcu, for $\sigma_{sd}^2=0$ dB, $\sigma_{sr}^2=15$ dB, $\sigma_{rd}^2=15$ dB, the achievable rate using DF scheme for 4-PSK and 8-PSK input constellations are about 3.5 dB and 1.5 dB away from that of the Gaussian alphabet. Hence in order to be within 1.5

dB away from the achievable rate of the DF scheme using Gaussian alphabet, for a spectral efficiency equal to 1 bpcu, 8 point signal set needs to be used. Hence, in the proposed TCM scheme, to achieve a spectral efficiency r bpcu, a signal set with 2^{2r+1} points is used.

Note 1: To achieve a spectral efficiency of r bpcu, a total of $2r$ bits get transmitted from S to D in the two Phases of relaying.

The value of E_S required for a BER of 10^{-4} and the diversity order for the different schemes achieving a spectral efficiency of 1 bpcu are shown in Table I, at the top of the next page (The details regarding the design of TCM schemes for the relay channel are presented in Section VI and E_S vs BER plots are given in Section VII. The TCM schemes for the direct transmission were designed based on the design criteria for TCM for fading channels [21]). From Table I, it can be seen that the proposed schemes for the relay channel, which exploit the fact that the S-R and R-D links are better than the S-D link, outperform the corresponding schemes for the direct transmission without relay. From Fig. 3, we see that the capacity of the direct transmission without relay with 4 PSK signal set and Gaussian alphabet are respectively 2 dB and 1.5 dB, which are greater than the value of E_S (0 dB) required for the 16 state TCM scheme for the relay channel, for a BER of 10^{-4} . This means that under the given assumptions, the proposed 16 state 8 PSK TCM scheme for the relay channel, will outperform even the best possible coding scheme for the direct transmission without relay. The claim made above is valid even for other values of the variances of the fading links, possibly with a higher number of states if not with 16 states.

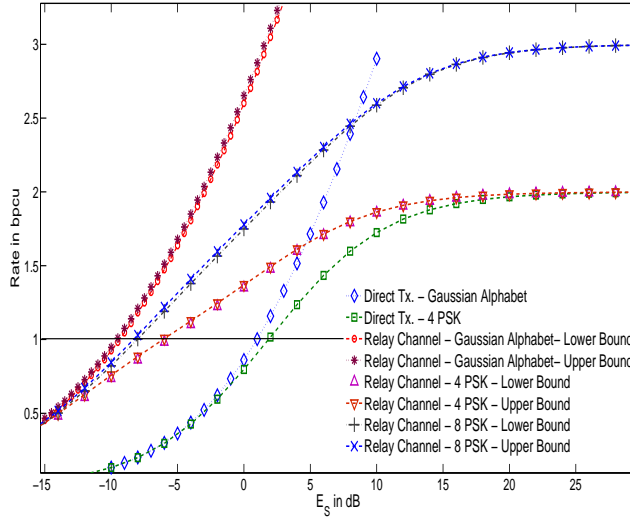


Fig. 3. Rate vs E_S for Half Duplex relay Channel with $\sigma_{sd}^2=0$ dB, $\sigma_{sr}^2=15$ dB, $\sigma_{rd}^2=15$ dB

III. A NEAR-ML TCM DECODER

The trellis used for decoding at D is constructed as described in the following subsection. The reason why the decoding trellis is of the form described will be clear by the end of Subsection B of this section.

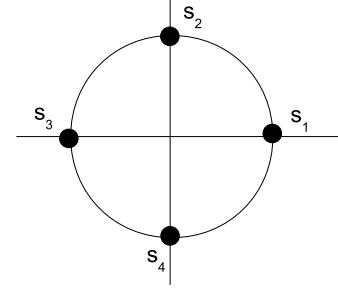


Fig. 4. 4-PSK signal set

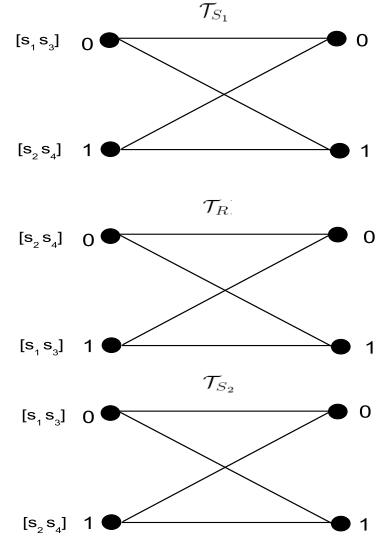


Fig. 5. Trellises \mathcal{T}_{S1} , \mathcal{T}_R and \mathcal{T}_{S2}

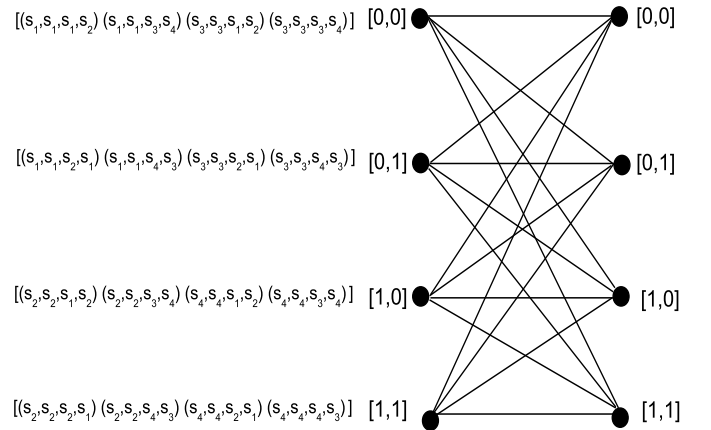


Fig. 6. Trellis \mathcal{T}_D used for decoding at D

TABLE I
COMPARISON OF DIFFERENT SCHEMES FOR A SPECTRAL EFFICIENCY OF 1 BPCU

Scheme	Diversity Order	E_S in dB for $BER = 10^{-4}$
Direct Tx. - BPSK	1	37 dB
Direct Tx. - 2 State TCM	2	18 dB
Direct Tx. - 4 State TCM	3	12 dB
Relay Channel - Uncoded Tx. Scheme	2	12.5 dB
Relay Channel - 2 State TCM	2	8 dB
Relay Channel - 4 State TCM	4	3 dB
Relay Channel - 8 State TCM	4	1.5 dB
Relay Channel - 16 State TCM	6	0 dB

A. Decoding Trellis at D

Let $\{a^i, a^{i+1}\}$ and $\{b^i, b^{i+1}\}$ respectively denote any two pairs of states in the i^{th} and $i + 1^{th}$ stages of the trellises \mathcal{T}_{S_1} and \mathcal{T}_R . Let $\mathcal{E}_{a^i, a^{i+1}}$ and $\mathcal{E}_{b^i, b^{i+1}}$ denote the sets of edges from a^i to a^{i+1} and b^i to b^{i+1} respectively.

The trellis used for decoding at D, denoted as \mathcal{T}_D , is constructed as follows: The states of \mathcal{T}_D at the i^{th} and $i + 1^{th}$ stages are denoted by the two tuples $[a^i, b^i]$ and $[a^{i+1}, b^{i+1}]$ respectively. The set of edges from $[a^i, b^i]$ to $[a^{i+1}, b^{i+1}]$, denoted as $\mathcal{E}_{[a^i, b^i], [a^{i+1}, b^{i+1}]}$, contains an edge denoted by the pair $\{e_{a^i, a^{i+1}}, e_{b^i, b^{i+1}}\}$ if $e_{a^i, a^{i+1}} \in \mathcal{E}_{a^i, a^{i+1}}$ and $e_{b^i, b^{i+1}} \in \mathcal{E}_{b^i, b^{i+1}}$. The edge $\{e_{a^i, a^{i+1}}, e_{b^i, b^{i+1}}\}$ is labelled with the four tuple $(\mathcal{X}_{s_1}(e_{a^i, a^{i+1}}), \mathcal{X}_{s_2}(e_{a^i, a^{i+1}}), \mathcal{X}_{s_1}(e_{b^i, b^{i+1}}), \mathcal{X}_r(e_{b^i, b^{i+1}})) \in S^4$.

Note 2: Even though encoding at S and R during the two Phases take place using a total of three trellises, the edges connecting the states in two successive stages in the trellis \mathcal{T}_D are constructed based on the trellises \mathcal{T}_{S_1} and \mathcal{T}_R . The reason is as follows: The uncoded bits at the input of the TCM encoders at S during Phase 1 and Phase 2 are the same. Hence, the encoded paths in the trellises \mathcal{T}_{S_1} and \mathcal{T}_{S_2} are always the same (though the complex signal sequences may be different, due to different labellings) and it is enough to consider one of the two trellises \mathcal{T}_{S_1} and \mathcal{T}_{S_2} .

Example 1: For the trellis triple $\{\mathcal{T}_{S_1}, \mathcal{T}_{S_2}, \mathcal{T}_R\}$ (shown in Fig. 5), labelled with signal points from S (shown in Fig. 4), the trellis \mathcal{T}_D is shown in Fig. 6.

The decoding metric for the proposed near-ML decoder is obtained in the following subsection.

B. Near-ML Decoding Metric

By assumption, R has perfect knowledge about the instantaneous value of the fade coefficient associated with the S-R link and D has perfect knowledge about the instantaneous values of the fade coefficients associated with the S-R, R-D and S-D links. Let \mathcal{A}_L denote the set of all paths of length L in the trellis (since the TCM encoder used at S during Phase 1, at S during Phase 2 and at R during Phase 2 are the same, the set \mathcal{A}_L is the same for the trellises \mathcal{T}_{S_1} , \mathcal{T}_{S_2} and \mathcal{T}_R). Let \mathcal{P}_S denote the path which corresponds to the symbols transmitted by the source during Phase 1. Let \mathcal{P}_R denote the path to which the Viterbi decoder at R decodes during Phase 1, i.e.,

$$\mathcal{P}_R = \arg \min_{\mathcal{P} \in \mathcal{A}_L} \sum_{i=1}^L |Y_r^i - h_{sr}^i x_{s_1}^i(\mathcal{P})|^2.$$

The optimal ML decoder at D decides in favour of the path,

$$\hat{\mathcal{P}}_S = \arg \max_{\mathcal{P}_S \in \mathcal{A}_L} Pr \left\{ Y_{d_1}^{1,L}, Y_{d_2}^{1,L} | \mathcal{P}_S \right\},$$

where $Y_{d_1}^{1,L}$ and $Y_{d_2}^{1,L}$ denote the sequences $Y_{d_1}^1, Y_{d_1}^2, \dots, Y_{d_1}^L$ and $Y_{d_2}^1, Y_{d_2}^2, \dots, Y_{d_2}^L$ respectively and $Pr \left\{ Y_{d_1}^{1,L}, Y_{d_2}^{1,L} | \mathcal{P}_S \right\}$ is the probability that $Y_{d_1}^{1,L}$ and $Y_{d_2}^{1,L}$ are the received sequences by D during Phase 1 and Phase 2, given that the path corresponding to the complex numbers transmitted by S is \mathcal{P}_S . The probability which needs to be maximized is given by,

$$Pr \left\{ Y_{d_1}^{1,L}, Y_{d_2}^{1,L} | \mathcal{P}_S \right\} = \sum_{\mathcal{P}_R \in \mathcal{A}_L} Pr \left\{ Y_{d_1}^{1,L}, Y_{d_2}^{1,L} | \mathcal{P}_S, \mathcal{P}_R \right\} Pr \left\{ \mathcal{P}_R | \mathcal{P}_S \right\}. \quad (10)$$

where $Pr \left\{ Y_{d_1}^{1,L}, Y_{d_2}^{1,L} | \mathcal{P}_S, \mathcal{P}_R \right\}$ is the probability that $Y_{d_1}^{1,L}$ and $Y_{d_2}^{1,L}$ are the received sequences by D during Phase 1 and Phase 2, given that the paths corresponding to the complex numbers transmitted by S and R are \mathcal{P}_S and \mathcal{P}_R respectively. The quantity $Pr \left\{ \mathcal{P}_R | \mathcal{P}_S \right\}$ is the probability that R decodes to the path \mathcal{P}_R given that \mathcal{P}_S is the path corresponding to the complex symbols transmitted by S.

The probability $Pr \left\{ \mathcal{P}_R | \mathcal{P}_S \right\}$ can be upper bounded by the corresponding PEP, i.e.,

$$\begin{aligned} Pr \left\{ \mathcal{P}_R | \mathcal{P}_S \right\} &\leq \prod_{i=1}^L Q \left[\frac{|h_{sr}^i (x_{s_1}^i(\mathcal{P}_R) - x_{s_1}^i(\mathcal{P}_S))|}{\sqrt{2}} \right] \\ &\leq \exp \left\{ -\frac{1}{4} \sum_{i=1}^L |h_{sr}^i (x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\mathcal{P}_R))|^2 \right\}, \quad (11) \end{aligned}$$

and we also have,

$$\begin{aligned} Pr \left\{ Y_{d_1}^{1,L}, Y_{d_2}^{1,L} | \mathcal{P}_S, \mathcal{P}_R \right\} &= \\ &\frac{1}{\pi^{2L}} \exp \left\{ -\sum_{i=1}^L \left(\left| Y_{d_1}^i - h_{sd_1}^i x_{s_1}^i(\mathcal{P}_S) \right|^2 \right. \right. \\ &\quad \left. \left. + \left| Y_{d_2}^i - h_{sd_2}^i x_{s_2}^i(\mathcal{P}_S) - h_{rd}^i x_r^i(\mathcal{P}_R) \right|^2 \right) \right\}. \quad (12) \end{aligned}$$

$$Pr\left\{Y_{d_1}^{1,L}, Y_{d_2}^{1,L} | \mathcal{P}_S\right\} \leq \frac{1}{\pi^{2L}} \sum_{\mathcal{P}_R \in \mathcal{A}_L} \exp \left\{ - \sum_{i=1}^L \left(\left| Y_{d_1}^i - h_{s_{d_1}}^i x_{s_1}^i(\mathcal{P}_S) \right|^2 + \left| Y_{d_2}^i - h_{s_{d_2}}^i x_{s_2}^i(\mathcal{P}_S) - h_{r_d}^i x_r^i(\mathcal{P}_R) \right|^2 + \frac{1}{4} \left| h_{sr}^i \left(x_{s_1}^i(\mathcal{P}_R) - x_{s_1}^i(\mathcal{P}_S) \right) \right|^2 \right) \right\} \quad (13)$$

$$\phi_i(\mathcal{P}_S, \mathcal{P}_R) = \left| Y_{d_1}^i - h_{s_{d_1}}^i x_{s_1}^i(\mathcal{P}_S) \right|^2 + \left| Y_{d_2}^i - h_{s_{d_2}}^i x_{s_2}^i(\mathcal{P}_S) - h_{r_d}^i x_r^i(\mathcal{P}_R) \right|^2 + \frac{1}{4} \left| h_{sr}^i \left(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\mathcal{P}_R) \right) \right|^2 \quad (14)$$

Substituting (11) and (12) in (10) we get (13), shown at the top of the next page.

If we maximize only the dominant exponential in the upper bound (13), then the decoded path is given by,

$$\hat{\mathcal{P}}_S = \arg \min_{\mathcal{P}_S \in \mathcal{A}_L} \left\{ \min_{\mathcal{P}_R \in \mathcal{A}_L} \sum_{i=1}^L \phi_i(\mathcal{P}_S, \mathcal{P}_R) \right\}, \quad (15)$$

where the metric $\phi_i(\mathcal{P}_S, \mathcal{P}_R)$ is given by (14) (shown at the top of the next page).

This near-ML decoder given by (15) involves minimizing the additive metric over two distinct paths and hence the decoding takes place in the trellis \mathcal{T}_D described in the previous subsection. Throughout, it is assumed that decoding at D takes place using this near-ML decoder.

The branch metric for the Viterbi decoder corresponding to the edge $\{e_{a^i, a^{i+1}}, e_{b^i, b^{i+1}}\}$ in \mathcal{T}_D can be obtained from (14) and is given by

$$\begin{aligned} f^i(e_{a^i, a^{i+1}}, e_{b^i, b^{i+1}}) &= \left| Y_{d_1}^i - h_{s_{d_1}}^i \mathcal{X}_{s_1}(e_{a^i, a^{i+1}}) \right|^2 \\ &\quad + \left| Y_{d_2}^i - h_{s_{d_2}}^i \mathcal{X}_{s_2}(e_{a^i, a^{i+1}}) - h_{r_d}^i \mathcal{X}_r(e_{b^i, b^{i+1}}) \right|^2 \\ &\quad + \frac{1}{4} \left| h_{sr}^i \left(\mathcal{X}_{s_1}(e_{a^i, a^{i+1}}) - \mathcal{X}_{s_1}(e_{b^i, b^{i+1}}) \right) \right|^2. \end{aligned}$$

For example, for the edge from state $[1, 0]$ to state $[0, 1]$ in \mathcal{T}_D in Example 1, denoted as $\{e_{1,0}, e_{0,1}\}$, from Fig. 6, the four tuple $(\mathcal{X}_{s_1}(e_{1,0}), \mathcal{X}_{s_2}(e_{1,0}), \mathcal{X}_{s_1}(e_{0,1}), \mathcal{X}_r(e_{0,1}))$ is (s_2, s_2, s_3, s_4) and hence the decoding metric is given by,

$$\begin{aligned} f^i(e_{1,0}, e_{0,1}) &= \left| Y_{d_1}^i - h_{s_{d_1}}^i s_2 \right|^2 + \left| Y_{d_2}^i - h_{s_{d_2}}^i s_2 - h_{r_d}^i s_4 \right|^2 \\ &\quad + \frac{1}{4} \left| h_{sr}^i (s_2 - s_3) \right|^2. \end{aligned}$$

Let K denote the number of branches which originate from each state and let N denote the number of states in the trellises \mathcal{T}_{S_1} , \mathcal{T}_{S_2} and \mathcal{T}_R . In the N^2 state trellis \mathcal{T}_D , K^2 branches originate from a state. As a result, in each step, the Viterbi algorithm used for decoding involves $N^2 K^2$ additions and N^2 comparisons of K^2 values. Hence the decoding complexity of the proposed near-ML decoder is approximately $2N^2 K^2 / (\log_2 K)$ operations per decoded bit, since $\log_2 K$ bits get transmitted in the two phases of relaying.

IV. PEP ANALYSIS OF THE NEAR-ML DECODER AND CODE DESIGN CRITERIA

The following Lemma is useful for the PEP analysis of the near-ML decoder.

Lemma 1: Consider a communication system with one transmitter and one receiver. The transmitter transmits one of the two n -dimensional complex vectors x_1 and $x_2 \in \mathbb{C}^n$ corresponding to two messages. The received n -dimensional complex vector $y = x + z \in \mathbb{C}^n$, where $x \in \{x_1, x_2\}$ and z is $CN(0, I_n)$. The decoding rule used at the receiver is as follows: The decoder decides in favour of message 1, if $\|y - x_1\|^2 \leq \|y - x_2\|^2 + c$ and it decides in favour of message 2 otherwise, where $c \in \mathbb{R}$ is a constant. Then the probability that the decoder decides in favour of message 2, given that message 1 was transmitted is upper bounded as,

$$Pr(1 \rightarrow 2) \leq \frac{1}{2} \exp \left\{ - \frac{\|x_1 - x_2\|^2}{4} - \frac{c}{2} \right\}.$$

Proof:

We have,

$$\begin{aligned} Pr(1 \rightarrow 2) &= Pr(\|y - x_1\|^2 \geq \|y - x_2\|^2 + c \mid x = x_1) \\ &= Pr\left(\Re\left\{\left(y - \frac{x_1 + x_2}{2}\right)^* (x_2 - x_1)\right\} \geq \frac{c}{2} \mid x = x_1\right) \\ &= Pr\left(\Re\left\{\left(z + \frac{x_1 - x_2}{2}\right)^* (x_2 - x_1)\right\} \geq \frac{c}{2}\right) \\ &= Pr\left(\Re\left\{z^* \frac{x_2 - x_1}{\|x_2 - x_1\|}\right\} \geq \frac{c}{2\|x_2 - x_1\|} + \frac{\|x_2 - x_1\|}{2}\right) \end{aligned}$$

Since z is $CN(0, I_n)$, it can be shown that $\Re\left\{z^* \frac{x_2 - x_1}{\|x_2 - x_1\|}\right\}$

is $N\left(0, \frac{1}{2}\right)$.

Hence,

$$\begin{aligned} Pr(1 \rightarrow 2) &= Q\left[\sqrt{2} \left(\frac{\|x_1 - x_2\|}{2} + \frac{c}{2\|x_2 - x_1\|} \right)\right] \\ &\leq \exp \left\{ - \left(\frac{\|x_1 - x_2\|}{2} + \frac{c}{2\|x_2 - x_1\|} \right)^2 \right\} \\ &\leq \exp \left\{ - \frac{\|x_1 - x_2\|^2}{4} - \frac{c}{2} \right\}. \end{aligned}$$

■

A. PEP Analysis of the Near-ML Decoder

In this subsection, an upper bound on the PEP that a transmitted path \mathcal{P}_S at S is decoded by the near-ML decoder at D as $\tilde{\mathcal{P}}_S$ is derived.

For the trellis \mathcal{T}_{S_1} , let $\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ denote the set of values of i for which $x_{s_1}^i(\mathcal{P}_S)$ and $x_{s_1}^i(\tilde{\mathcal{P}}_S)$ are different, for $1 \leq$

$i \leq L$. Similarly, $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ and $\eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ are defined for the trellises \mathcal{T}_{S_2} and \mathcal{T}_R respectively.

Theorem 1: For the proposed TCM scheme, the PEP that a transmitted path \mathcal{P}_S in the trellis is decoded as $\tilde{\mathcal{P}}_S$, under the proposed near-ML decoder, is upper bounded by (16) (shown at the top of the next page).

Proof: The PEP that a transmitted path \mathcal{P}_S is decoded as $\tilde{\mathcal{P}}_S$ at D is

$$Pr\{\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S\} = \sum_{\mathcal{P}_R \in \mathcal{A}_L} Pr\{\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S | \mathcal{P}_R\} Pr\{\mathcal{P}_R | \mathcal{P}_S\}, \quad (17)$$

where \mathcal{P}_R is the path to which R decodes. The probability $Pr\{\mathcal{P}_R | \mathcal{P}_S\}$ is upper bounded by (11).

Also,

$$\begin{aligned} & Pr\{\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S | \mathcal{P}_R\} \\ &= Pr\left\{ \min_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \sum_{i=1}^L \phi_i(\mathcal{P}_S, \tilde{\mathcal{P}}_R) \geq \min_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \sum_{i=1}^L \phi_i(\tilde{\mathcal{P}}_S, \tilde{\mathcal{P}}_R) \mid (\mathcal{P}_S, \mathcal{P}_R) \right\}. \end{aligned} \quad (18)$$

Let \mathcal{F}_1 and \mathcal{F}_2 be two sets defined as,

$$\begin{aligned} \mathcal{F}_1 &= \left\{ (Y_{d_1}^{1,L}, Y_{d_2}^{1,L}) \mid \min_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \sum_{i=1}^L \phi_i(\mathcal{P}_S, \tilde{\mathcal{P}}_R) \right. \\ &\quad \left. \geq \min_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \sum_{i=1}^L \phi_i(\tilde{\mathcal{P}}_S, \tilde{\mathcal{P}}_R) \right\}, \\ \mathcal{F}_2 &= \left\{ (Y_{d_1}^{1,L}, Y_{d_2}^{1,L}) \mid \sum_{i=1}^L \phi_i(\mathcal{P}_S, \mathcal{P}_R) \geq \min_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \sum_{i=1}^L \phi_i(\tilde{\mathcal{P}}_S, \tilde{\mathcal{P}}_R) \right\}. \end{aligned}$$

Since $\mathcal{F}_1 \subset \mathcal{F}_2$, from (18) we have,

$$\begin{aligned} & Pr\{\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S | \mathcal{P}_R\} \\ &\leq Pr\left\{ \sum_{i=1}^L \phi_i(\mathcal{P}_S, \mathcal{P}_R) \geq \min_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \sum_{i=1}^L \phi_i(\tilde{\mathcal{P}}_S, \tilde{\mathcal{P}}_R) \mid (\mathcal{P}_S, \mathcal{P}_R) \right\} \end{aligned} \quad (19)$$

Let \mathcal{F}_3 be the set defined as,

$$\mathcal{F}_3 = \bigcup_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \left\{ (Y_{d_1}^{1,L}, Y_{d_2}^{1,L}) \mid \sum_{i=1}^L \phi_i(\mathcal{P}_S, \mathcal{P}_R) \geq \sum_{i=1}^L \phi_i(\tilde{\mathcal{P}}_S, \tilde{\mathcal{P}}_R) \right\}.$$

From (19), noting that $\mathcal{F}_2 \subset \mathcal{F}_3$ and using the union bound, we have,

$$\begin{aligned} & Pr\{\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S | \mathcal{P}_R\} \\ &\leq \sum_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} Pr\left\{ \sum_{i=1}^L \phi_i(\mathcal{P}_S, \mathcal{P}_R) \geq \sum_{i=1}^L \phi_i(\tilde{\mathcal{P}}_S, \tilde{\mathcal{P}}_R) \mid (\mathcal{P}_S, \mathcal{P}_R) \right\}. \end{aligned} \quad (20)$$

The probability inside the summation in (20) can be upper bounded using Lemma 1. The constant c , the vectors x_i , $i \in \{1, 2\}$ and the vector y , defined in Lemma 1, are chosen as given by (21)-(24) (shown at the top of the next page).

Using Lemma 1 in (20) gives rise to (25) ((25)-(27) are shown on the next page). Substituting (11) and (25) in (17) gives (26). Taking expectation in (26) with respect to $h_{sd_1}^i$, $h_{sd_2}^i$, h_{rd}^i and h_{sr}^i , $1 \leq i \leq L$, yields (27) which is the same as (16). \blacksquare

The criteria to maximize the diversity order of the proposed TCM scheme is obtained in the following subsection.

B. Diversity Criteria

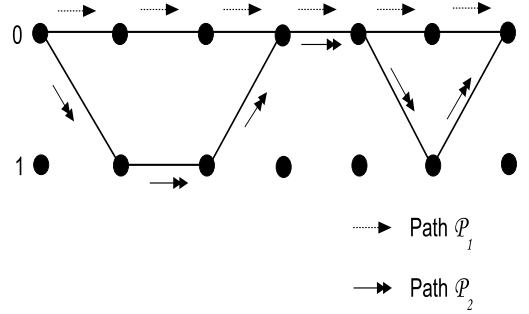


Fig. 7. Example illustrating the notion of unmerged length between two paths

Definition 1: The *unmerged length* between the paths \mathcal{P}_1 and \mathcal{P}_2 in the trellis, denoted as $h(\mathcal{P}_1, \mathcal{P}_2)$ is the number of branches in which the paths \mathcal{P}_1 and \mathcal{P}_2 differ. For example, for the paths \mathcal{P}_1 and \mathcal{P}_2 of length 6 shown in Fig. 7, the unmerged length $h(\mathcal{P}_1, \mathcal{P}_2) = 5$. The *unmerged length* of the code is the minimum value of $h(\mathcal{P}_1, \mathcal{P}_2)$, over all possible pairs of paths $(\mathcal{P}_1, \mathcal{P}_2)$ in the trellis.

Note 3: Since the unmerged length between a pair of paths $\mathcal{P}_1, \mathcal{P}_2$ is the same in all three trellises \mathcal{T}_{S_1} , \mathcal{T}_{S_2} and \mathcal{T}_R , we simply refer to it as the unmerged length in the trellis, without mentioning whether the trellis is \mathcal{T}_{S_1} , \mathcal{T}_{S_2} or \mathcal{T}_R .

For two paths \mathcal{P}_S and $\tilde{\mathcal{P}}_S$ whose unmerged length is equal to the unmerged length of the code, the first term in the PEP upper bound given by (16) is of diversity order upper bounded by twice the unmerged length of the code, since both $|\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$ and $|\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$ are upper bounded by the unmerged length of the code. Hence, the diversity order of the proposed TCM scheme is upper bounded by twice the unmerged length of the code. The mappings \mathcal{X}_{S_1} , \mathcal{X}_{S_2} and \mathcal{X}_R should be chosen such that this bound is met with equality.

For two arbitrary paths \mathcal{P}_S and $\tilde{\mathcal{P}}_S$ in the trellis, let the diversity order corresponding to the error event $(\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S)$ be denoted as $D(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$.

Corollary 1: For the proposed TCM scheme, $D(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is bounded as,

$$\begin{aligned}
Pr\left\{\mathcal{P}_S \longrightarrow \tilde{\mathcal{P}}_S\right\} \leq & \left(\prod_{i \in \eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)} \left[\frac{1}{1 + \frac{|\sigma_{sd}(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\tilde{\mathcal{P}}_S))|^2}{4}} \right] \right. \\
& \left. \prod_{i \in \{\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)\}} \left[\frac{1}{1 + \frac{|\sigma_{sd}(x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S))|^2 + |\sigma_{rd}(x_r^i(\mathcal{P}_S) - x_r^i(\tilde{\mathcal{P}}_S))|^2}{4}} \right] \right) \\
& + \sum_{\substack{\mathcal{P}_R \in \mathcal{A}_L \\ \mathcal{P}_R \neq \mathcal{P}_S}} \sum_{\substack{\tilde{\mathcal{P}}_R \in \mathcal{A}_L \\ \tilde{\mathcal{P}}_R \neq \tilde{\mathcal{P}}_S}} \left(\prod_{i \in \eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)} \left[\frac{1}{1 + \frac{|\sigma_{sd}(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\tilde{\mathcal{P}}_S))|^2}{4}} \right] \right. \\
& \prod_{i \in \{\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)\}} \left[\frac{1}{1 + \frac{|\sigma_{sd}(x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S))|^2 + |\sigma_{rd}(x_r^i(\mathcal{P}_R) - x_r^i(\tilde{\mathcal{P}}_R))|^2}{4}} \right] \\
& \left. \prod_{i \in \{\eta_{s_1}(\mathcal{P}_S, \mathcal{P}_R) \cup \eta_{s_1}(\tilde{\mathcal{P}}_S, \tilde{\mathcal{P}}_R)\}} \left[\frac{1}{1 + \frac{|\sigma_{sr}(x_{s_1}^i(\tilde{\mathcal{P}}_R) - x_{s_1}^i(\tilde{\mathcal{P}}_S))|^2 + |\sigma_{sr}(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\mathcal{P}_R))|^2}{8}} \right] \right) \quad (16)
\end{aligned}$$

$$c = \frac{1}{4} \sum_{i=1}^L \left(\left| h_{sr}^i(x_{s_1}^i(\tilde{\mathcal{P}}_R) - x_{s_1}^i(\tilde{\mathcal{P}}_S)) \right|^2 - \left| h_{sr}^i(x_{s_1}^i(\mathcal{P}_R) - x_{s_1}^i(\mathcal{P}_S)) \right|^2 \right) \quad (21)$$

$$x_1 = [h_{sd_1}^1 x_{s_1}^1(\mathcal{P}_S), h_{sd_1}^2 x_{s_1}^2(\mathcal{P}_S), \dots, h_{sd_1}^L x_{s_1}^L(\mathcal{P}_S), h_{sd_2}^1 x_{s_2}^1(\mathcal{P}_S) + h_{rd}^1 x_{rd}^1(\mathcal{P}_R), h_{sd_2}^2 x_{s_2}^2(\mathcal{P}_S) + h_{rd}^2 x_{rd}^2(\mathcal{P}_R), \dots, h_{sd_2}^L x_{s_2}^L(\mathcal{P}_S) + h_{rd}^L x_{rd}^L(\mathcal{P}_R)]^T \quad (22)$$

$$x_2 = [h_{sd_1}^1 x_{s_1}^1(\tilde{\mathcal{P}}_S), h_{sd_1}^2 x_{s_1}^2(\tilde{\mathcal{P}}_S), \dots, h_{sd_1}^L x_{s_1}^L(\tilde{\mathcal{P}}_S), h_{sd_2}^1 x_{s_2}^1(\tilde{\mathcal{P}}_S) + h_{rd}^1 x_{rd}^1(\tilde{\mathcal{P}}_R), h_{sd_2}^2 x_{s_2}^2(\tilde{\mathcal{P}}_S) + h_{rd}^2 x_{rd}^2(\tilde{\mathcal{P}}_R), \dots, h_{sd_2}^L x_{s_2}^L(\tilde{\mathcal{P}}_S) + h_{rd}^L x_{rd}^L(\tilde{\mathcal{P}}_R)]^T \quad (23)$$

$$y = [Y_{d_1}^1, Y_{d_1}^2, \dots, Y_{d_1}^L, Y_{d_2}^1, Y_{d_2}^2, \dots, Y_{d_2}^L]^T \quad (24)$$

$$\begin{aligned}
Pr\left\{\mathcal{P}_S \longrightarrow \tilde{\mathcal{P}}_S | \mathcal{P}_R\right\} \leq & \sum_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \exp \left\{ \sum_{i=1}^L \left(-\frac{1}{2} \left| h_{sd_1}^i(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\tilde{\mathcal{P}}_S)) \right|^2 - \frac{1}{2} \left| h_{sd_2}^i(x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S)) + h_{rd}^i(x_r^i(\mathcal{P}_R) - x_r^i(\tilde{\mathcal{P}}_R)) \right|^2 \right. \right. \\
& \left. \left. - \frac{1}{8} \left| h_{sr}^i(x_{s_1}^i(\tilde{\mathcal{P}}_R) - x_{s_1}^i(\tilde{\mathcal{P}}_S)) \right|^2 + \frac{1}{8} \left| h_{sr}^i(x_{s_1}^i(\mathcal{P}_R) - x_{s_1}^i(\mathcal{P}_S)) \right|^2 \right) \right\} \quad (25)
\end{aligned}$$

$$\begin{aligned}
Pr\left\{\mathcal{P}_S \longrightarrow \tilde{\mathcal{P}}_S\right\} \leq & \sum_{\mathcal{P}_R \in \mathcal{A}_L} \sum_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \exp \left\{ \sum_{i=1}^L \left(-\frac{1}{2} \left| h_{sd_1}^i(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\tilde{\mathcal{P}}_S)) \right|^2 - \frac{1}{2} \left| h_{sd_2}^i(x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S)) + h_{rd}^i(x_r^i(\mathcal{P}_R) - x_r^i(\tilde{\mathcal{P}}_R)) \right|^2 \right. \right. \\
& \left. \left. - \frac{1}{8} \left| h_{sr}^i(x_{s_1}^i(\tilde{\mathcal{P}}_R) - x_{s_1}^i(\tilde{\mathcal{P}}_S)) \right|^2 - \frac{1}{8} \left| h_{sr}^i(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\mathcal{P}_R)) \right|^2 \right) \right\} \quad (26)
\end{aligned}$$

$$\begin{aligned}
Pr\left\{\mathcal{P}_S \longrightarrow \tilde{\mathcal{P}}_S\right\} \leq & \sum_{\mathcal{P}_R \in \mathcal{A}_L} \sum_{\tilde{\mathcal{P}}_R \in \mathcal{A}_L} \prod_{i=1}^L \left(\left[\frac{1}{1 + \frac{|\sigma_{sd}(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\tilde{\mathcal{P}}_S))|^2}{4}} \right] \left[\frac{1}{1 + \frac{|\sigma_{sd}(x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S))|^2 + |\sigma_{rd}(x_r^i(\mathcal{P}_R) - x_r^i(\tilde{\mathcal{P}}_R))|^2}{4}} \right] \right. \\
& \left. \left[\frac{1}{1 + \frac{|\sigma_{sr}(x_{s_1}^i(\tilde{\mathcal{P}}_R) - x_{s_1}^i(\tilde{\mathcal{P}}_S))|^2 + |\sigma_{sr}(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\mathcal{P}_R))|^2}{8}} \right] \right) \quad (27)
\end{aligned}$$

$$Pr\left(\mathcal{P}_S \longrightarrow \tilde{\mathcal{P}}_S\right) \leq K \prod_{i \in \eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)} \left[\frac{1}{|\sigma_{sd}(x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\tilde{\mathcal{P}}_S))|^2} \right] \prod_{i \in \{\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)\}} \left[\frac{1}{|\sigma_{sd}(x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S))|^2 + |\sigma_{rd}(x_r^i(\mathcal{P}_S) - x_r^i(\tilde{\mathcal{P}}_S))|^2} \right] \quad (28)$$

$$\begin{aligned}
|\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| + |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| &\leq D(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \\
&\leq \min \left\{ |\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| + |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|, \right. \\
&\quad \left. 2|\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| + |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| \right\}, \tag{29}
\end{aligned}$$

with equality on both the sides if and only if $\eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is a subset of $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$.

Proof: From (16), $D(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is given by,

$$\begin{aligned}
D(\mathcal{P}_S, \tilde{\mathcal{P}}_S) &= |\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| + \min_{\mathcal{P}_R, \tilde{\mathcal{P}}_R} \left\{ |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)| \right. \\
&\quad \left. + |\eta_{s_1}(\tilde{\mathcal{P}}_R, \tilde{\mathcal{P}}_S) \cup \eta_{s_1}(\mathcal{P}_R, \tilde{\mathcal{P}}_S)| \right\}, \\
&\geq |\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| + \min_{\mathcal{P}_R, \tilde{\mathcal{P}}_R} |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)| \\
&\quad + \min_{\mathcal{P}_R, \tilde{\mathcal{P}}_R} |\eta_{s_1}(\tilde{\mathcal{P}}_R, \tilde{\mathcal{P}}_S) \cup \eta_{s_1}(\mathcal{P}_R, \tilde{\mathcal{P}}_S)|. \tag{30}
\end{aligned}$$

The quantity $|\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)|$ attains the minimum value, for the case when $\mathcal{P}_R = \tilde{\mathcal{P}}_R$, in which case $\eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)$ is an empty set, or for the case when the set $\eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)$ is a subset of $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$. In both the cases the minimum value attained is $|\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$.

The minimum value of $|\eta_{s_1}(\tilde{\mathcal{P}}_R, \tilde{\mathcal{P}}_S) \cup \eta_{s_1}(\mathcal{P}_R, \tilde{\mathcal{P}}_S)|$ equal to zero is attained when $\mathcal{P}_R = \mathcal{P}_S$ and $\tilde{\mathcal{P}}_R = \tilde{\mathcal{P}}_S$.

Hence from (30) we get the lower bound in (29).

The lower bound in (29) is met with equality if both the minima in (30) are attained for the same choice of the pair $(\mathcal{P}_R, \tilde{\mathcal{P}}_R)$, which occurs if and only if $\mathcal{P}_R = \mathcal{P}_S$, $\tilde{\mathcal{P}}_R = \tilde{\mathcal{P}}_S$ and the set $\eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)$ is a subset of $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$. In other words, equality in (30) occurs if and only if $\eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is a subset of $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$.

The choice $\mathcal{P}_R = \mathcal{P}_S$ and $\tilde{\mathcal{P}}_R = \tilde{\mathcal{P}}_S$ results in,

$$\begin{aligned}
&|\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)| \\
&+ |\eta_{s_1}(\tilde{\mathcal{P}}_R, \tilde{\mathcal{P}}_S) \cup \eta_{s_1}(\mathcal{P}_R, \tilde{\mathcal{P}}_S)| = |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|. \tag{31}
\end{aligned}$$

The choice $\mathcal{P}_R = \tilde{\mathcal{P}}_R = \tilde{\mathcal{P}}_S$ results in,

$$\begin{aligned}
&|\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_R, \tilde{\mathcal{P}}_R)| \\
&+ |\eta_{s_1}(\tilde{\mathcal{P}}_R, \tilde{\mathcal{P}}_S) \cup \eta_{s_1}(\mathcal{P}_R, \tilde{\mathcal{P}}_S)| = |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| + |\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|. \tag{32}
\end{aligned}$$

From (31) and (32) we get the upper bound in (29).

Equality in both the lower and upper bounds of (29) occurs if and only if $\eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is a subset of $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$. ■

Definition 2: The *effective length* of a pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis \mathcal{T}_{S_1} is defined to be the cardinality of the set $\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$. The *effective length* of the trellis \mathcal{T}_{S_1} is defined to be minimum among the effective lengths of all possible pairs of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$. In a similar way, the effective lengths of the trellises \mathcal{T}_{S_2} and \mathcal{T}_R can be defined.

Definition 3: The *generalized effective length* of a pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis pair $(\mathcal{T}_{S_2}, \mathcal{T}_R)$ is defined to be the cardinality of the set $\{\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)\}$.

Note 4: If the labellings \mathcal{X}_{s_2} and \mathcal{X}_r are the same, the generalized effective length of the pair of paths reduces to the effective length of the pair of paths in the trellis \mathcal{T}_{S_2} , which is equal to the effective length of the pair of paths in the trellis \mathcal{T}_R .

If \mathcal{X}_{s_1} and \mathcal{X}_{s_2} are chosen such that the effective lengths of the trellises \mathcal{T}_{S_1} and \mathcal{T}_{S_2} are equal to the unmerged length of the code, then for every pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, $|\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$ and $|\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$ are greater than the unmerged length of the code. In that case, from the lower bound in Corollary 1 and from the fact that the diversity order cannot exceed twice the unmerged length of the code, it follows that the diversity order for the proposed scheme is equal to twice the unmerged length of the code. Hence, *a sufficient condition to obtain maximum diversity order is to maximize the effective lengths of the trellises \mathcal{T}_{S_1} and \mathcal{T}_{S_2}* , independent of the labelling \mathcal{X}_r . This is not surprising, since during Phase 2 the source encodes the same bits it encoded and transmitted during Phase 1, maximum diversity order is obtained irrespective of whether the relay transmits during Phase 2 or not. But since $\sigma_{sd}^2 \ll \sigma_{rd}^2$, from (28) it can be seen that the coding gain is greatly reduced if the relay does not transmit during Phase 2. It is shown in the following subsection that the effective length of \mathcal{T}_R also should equal the unmerged length of the code to avoid this reduction in coding gain.

Even though the above condition guarantees maximum diversity, there is no simple expression for $D(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ (Corollary 1 gives only the upper and lower bounds) and hence it is not clear which error events contribute to the minimum diversity order. Proposition 1 below gives the conditions under which an error event results in the minimum diversity order.

Proposition 1: Assuming that the sufficient condition to obtain full diversity is satisfied, the error event corresponding to a pair of paths contributes to the minimum diversity order if and only if the effective length of the pair of paths in \mathcal{T}_{S_1} and the generalized effective length of the pair of paths in $(\mathcal{T}_{S_2}, \mathcal{T}_R)$, are equal to the unmerged length of the code.

Proof: Let \mathcal{P}_S and $\tilde{\mathcal{P}}_S$ be two paths in the trellis. The maps \mathcal{X}_{S_1} and \mathcal{X}_{S_2} are chosen such that the effective lengths of the trellises \mathcal{T}_{S_1} and \mathcal{T}_{S_2} are equal to the unmerged length of the code.

From the lower bound in Corollary 1, $D(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is greater than or equal to twice the unmerged length of the code, since $|\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$ and $|\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$ are greater than or equal to the unmerged length of the code. Equality can occur only when the set $\eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is a subset of $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, and the effective lengths of the pair $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in both the trellises \mathcal{T}_{S_1} and \mathcal{T}_{S_2} , i.e., $|\eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$ and $|\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$, are equal to the unmerged length of the code. Clearly, $\eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is not a subset of $\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, if and only if $|\eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)| > |\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)|$, in which case the generalized effective length of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis pair $(\mathcal{T}_{S_2}, \mathcal{T}_R)$ is greater than unmerged length of the code. Hence, the error event corresponding to a pair of paths give rise to the minimum diversity order, if and only if the conditions given in the statement of the Proposition are satisfied. ■

In the rest of the paper, it is assumed that the sufficient condition to obtain maximum diversity is satisfied.

For a pair of paths \mathcal{P}_S and $\tilde{\mathcal{P}}_S$ resulting in the minimum diversity, only the first term in (16) gives rise to minimum diversity order. Neglecting terms of higher diversity order in (16), we get (28) (shown on the previous to the previous page), where K is a positive constant.

The criterion to maximize the coding gain is obtained in the following subsection.

C. Coding Gain Criterion

Let $m_1(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ denote the product distance of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis \mathcal{T}_{S_1} , i.e.,

$$m_1(\mathcal{P}_S, \tilde{\mathcal{P}}_S) = \prod_{i \in \eta_{s_1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)} \left| x_{s_1}^i(\mathcal{P}_S) - x_{s_1}^i(\tilde{\mathcal{P}}_S) \right|^2.$$

Definition 4: The *generalized product distance* of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ corresponding to the trellis pair $(\mathcal{T}_{S_2}, \mathcal{T}_R)$ denoted as $m_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is defined as,

$$m_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S) = \prod_{i \in \{\eta_{s_2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)\}} \left(\gamma \left| x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S) \right|^2 + \left| x_r^i(\mathcal{P}_S) - x_r^i(\tilde{\mathcal{P}}_S) \right|^2 \right), \quad (33)$$

where $\gamma = \sigma_{sd}^2 / \sigma_{rd}^2$.

Note 5: If the labelling used for the trellises \mathcal{T}_R and \mathcal{T}_{S_2} are the same, the generalized product distance of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ corresponding to the trellis pair $(\mathcal{T}_{S_2}, \mathcal{T}_R)$, reduces within a constant scaling factor to the product distance of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis \mathcal{T}_R , which is equal to the product distance of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis \mathcal{T}_{S_2} .

Definition 5: Consider the trellis triplet $(\mathcal{T}_{S_1}, \mathcal{T}_{S_2}, \mathcal{T}_R)$. The combined product distance of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, denoted by $m(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, is defined as the product of the product distance of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis \mathcal{T}_{S_1} and the generalised product distance of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ corresponding to the trellis pair $(\mathcal{T}_{S_2}, \mathcal{T}_R)$, i.e.,

$$m(\mathcal{P}_S, \tilde{\mathcal{P}}_S) = \left[m_1(\mathcal{P}_S, \tilde{\mathcal{P}}_S) m_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \right].$$

Definition 6: The combined product distance of the trellis triplet $(\mathcal{T}_{S_1}, \mathcal{T}_{S_2}, \mathcal{T}_R)$, which is also the coding gain metric denoted by \mathcal{G} , is defined to be the minimum value of $m(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, i.e.,

$$\mathcal{G} = \min_{(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \in \mathcal{Z}_{\mathcal{P}}} \left[m(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \right].$$

where $\mathcal{Z}_{\mathcal{P}}$ denotes the set of all pairs of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ whose effective length in \mathcal{T}_{S_1} and the generalized effective length in $(\mathcal{T}_{S_2}, \mathcal{T}_R)$ are equal to the unmerged length of the code.

From (28), it can be seen that *the combined product distance of the trellis triplet $(\mathcal{T}_{S_1}, \mathcal{T}_{S_2}, \mathcal{T}_R)$ needs to be maximized, to maximize the coding gain.* For a pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ which

contribute to the minimum diversity order, if the effective length of $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ in the trellis \mathcal{T}_R is not equal to the unmerged length of the code, then $x_r^i(\mathcal{P}_S) = x_r^i(\tilde{\mathcal{P}}_S)$ for some $i \in \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$. In that case $m_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ is greatly reduced since $\gamma \ll 1$. In order to avoid this, the effective length of the trellis \mathcal{T}_R should be made equal to the unmerged length of the code.

D. Near Optimality of the Proposed Near-ML decoder

The following argument proves the high SNR near optimality of the proposed near-ML decoder for the TCM scheme. Consider the situation where the S-R link is ideal, i.e., the relay decodes all the bits correctly. The optimal ML decoder at D decides in favour of the path given by,

$$\hat{\mathcal{P}}_S = \arg \min_{\mathcal{P}_S \in \mathcal{A}_L} \left(\sum_{i=1}^L \left| Y_{d_1}^i - h_{sd_1}^i x_{s_1}^i(\mathcal{P}_S) \right|^2 + \left| Y_{d_2}^i - h_{sd_2}^i x_{s_2}^i(\mathcal{P}_S) - h_{rd}^i x_r^i(\mathcal{P}_S) \right|^2 \right).$$

For this case, the PEP that the path \mathcal{P}_S is decoded as $\tilde{\mathcal{P}}_S$ is upper bounded by (34) (shown at the top of the next page). Taking expectation with respect to the fading coefficients, we get (35) (shown at the top of the next page). From (28) and (35), we see that the high SNR bounds on the PEP for the proposed near ML decoder with a non-ideal S-R link is same as that of the optimal ML decoder with an ideal S-R link, if we consider only the error events giving rise to the minimum diversity order. The simulation results presented in Section VII confirm that the BER vs SNR performance of the proposed near ML decoder with a non-ideal S-R link approaches the performance of optimal ML decoder with an ideal S-R link, at high SNR. The high SNR performance of the optimal ML decoder for the proposed TCM scheme, with a non-ideal S-R link cannot be better than that of optimal ML decoder for the case when the S-R link is ideal. This implies the high SNR performance of the proposed near-ML decoder approaches the performance of the optimal ML decoder.

From (28), it is clearly seen that even for the uncoded transmission scheme, choosing the labelling schemes used at S and R properly can potentially yield a significant performance improvement. The effect of the choice of labelling on the performance for the uncoded transmission scheme is discussed in the following section.

V. THE UNCODED TRANSMISSION SCHEME

Consider the uncoded transmission scheme in which bits are directly mapped onto complex symbols at S and R. A collection of $\log_2 M$ bits constitutes a message. Let $\mathcal{M} = \{1, 2, \dots, M\}$ denote this message set. The uncoded transmission scheme has an equivalent one state trellis representation, with M edges connecting two successive stages. The unmerged length between every pair of distinct paths is one and hence no distinction needs to be made between paths and edges. The index i in the transmitted complex numbers $x_{s_1}^i(\cdot)$, $x_{s_2}^i(\cdot)$ and $x_r^i(\cdot)$ can be dropped and they are the same as the

$$Pr\{\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S\} \leq \exp\left\{\sum_{i=1}^L \left(-\frac{1}{2} \left|h_{sd1}^i(x_{s1}^i(\mathcal{P}_S) - x_{s1}^i(\tilde{\mathcal{P}}_S))\right|^2 - \frac{1}{2} \left|h_{sd2}^i(x_{s2}^i(\mathcal{P}_S) - x_{s2}^i(\tilde{\mathcal{P}}_S)) + h_{rd}^i(x_r^i(\mathcal{P}_S) - x_r^i(\tilde{\mathcal{P}}_S))\right|^2\right)\right\} \quad (34)$$

$$Pr(\mathcal{P}_S \rightarrow \tilde{\mathcal{P}}_S) \leq K \prod_{i \in \eta_{s1}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)} \left[\frac{1}{|\sigma_{sd}(x_{s1}^i(\mathcal{P}_S) - x_{s1}^i(\tilde{\mathcal{P}}_S))|^2} \right] \prod_{i \in \{\eta_{s2}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \cup \eta_r(\mathcal{P}_S, \tilde{\mathcal{P}}_S)\}} \left[\frac{1}{|\sigma_{sd}(x_{s2}^i(\mathcal{P}_S) - x_{s2}^i(\tilde{\mathcal{P}}_S))|^2 + |\sigma_{rd}(x_r^i(\mathcal{P}_S) - x_r^i(\tilde{\mathcal{P}}_S))|^2} \right] \quad (35)$$

labellings on the edges $\mathcal{X}_{s1}(\cdot)$, $\mathcal{X}_{s2}(\cdot)$ and $\mathcal{X}_r(\cdot)$. Each path (edge) in the trellis is identified by the message $a \in \mathcal{M}$ which gets transmitted.

Corollary 2: For the uncoded transmission scheme, the PEP that the decoder at D decides in favour of message $\bar{a} \in \mathcal{M}$ given that the message transmitted by the source was $a \in \mathcal{M}$ is upper bounded as,

$$Pr(a \rightarrow \bar{a}) \leq \left[\frac{1}{1 + \frac{1}{4}|\sigma_{sd}|^2|x_{s1}(a) - x_{s1}(\bar{a})|^2} \right] \left[\frac{1}{1 + \frac{1}{4}|\sigma_{sd}|^2|x_{s2}(a) - x_{s2}(\bar{a})|^2 + \frac{1}{4}|\sigma_{rd}|^2|x_r(a) - x_r(\bar{a})|^2} \right] + H.O.T.$$

where H.O.T denotes the terms of diversity order greater than 2.

Proof: From (16), replacing the paths \mathcal{P}_S , $\tilde{\mathcal{P}}_S$, \mathcal{P}_R and $\tilde{\mathcal{P}}_R$ by the corresponding message indices, we get (36) (shown at the top of the next page). Neglecting terms of diversity order greater than 2 gives the result. ■

From Corollary 2, it follows that the diversity order of the uncoded transmission scheme is two and in order to minimise the PEP, we need to maximize the following:

$$\left[1 + \frac{1}{4}|\sigma_{sd}|^2|x_{s1}(a) - x_{s1}(\bar{a})|^2 \right] \left[1 + \frac{1}{4}|\sigma_{sd}|^2|x_{s2}(a) - x_{s2}(\bar{a})|^2 + \frac{1}{4}|\sigma_{rd}|^2|x_r(a) - x_r(\bar{a})|^2 \right].$$

At high SNR we need to maximize the metric,

$$m(a, \bar{a}) = |x_{s1}(a) - x_{s1}(\bar{a})|^2 [\gamma|x_{s2}(a) - x_{s2}(\bar{a})|^2 + |x_r(a) - x_r(\bar{a})|^2],$$

over all message pairs (a, \bar{a}) .

By a labelling scheme, we refer to the triplet $(\mathcal{X}_{s1}, \mathcal{X}_{s2}, \mathcal{X}_r)$. Let \mathcal{L} denote a labelling scheme used at S and R. Let us define,

$$d(\mathcal{L}) = \min_{a, \bar{a}, a \neq \bar{a}} m(a, \bar{a}),$$

where $a, \bar{a} \in \mathcal{M}$.

Let \mathcal{L}_0 denote the labelling scheme in which the mapping from bits to complex symbols, used by S (during Phase 1 and

Phase 2) and R (during Phase 2), are the same. Similar to $d(\mathcal{L})$, $d(\mathcal{L}_0)$ can be defined for the labelling scheme \mathcal{L}_0 .

Definition 7: The Labelling Gain of the labelling scheme \mathcal{L} , which is a measure of the performance gain provided by \mathcal{L} over \mathcal{L}_0 , is given by,

$$L_G(\mathcal{L}) = 10 \log_{10} \left[\frac{d(\mathcal{L})}{d(\mathcal{L}_0)} \right] \text{ dB}.$$

It is important to note that the Labelling Gain is calculated based on the upper bound on the PEP, taking into consideration only those pair of messages a and \bar{a} which contribute dominantly to the metric $m(a, \bar{a})$. The actual high SNR gain provided by the labelling scheme \mathcal{L} over the scheme \mathcal{L}_0 need not equal $L_G(\mathcal{L})$.

Throughout, the phrase *with our labelling* means that S and R use the labelling scheme which is described in the following subsection and *with constant labelling* means that S and R use the labelling scheme \mathcal{L}_0 .

A. A Labelling Scheme for PSK Constellation

In this subsection, we provide a good labelling scheme for 2^l -PSK constellation, where $l \geq 2$ is an integer. Let $s_k = \exp(jk2\pi/M)$, where $0 \leq k \leq M-1$ and $M = 2^l$, denote the signal points in the 2^l -PSK signal set.

Let $\tilde{\mathcal{L}}$ denote the labelling scheme described as follows: The maps \mathcal{X}_{s1} , \mathcal{X}_r and \mathcal{X}_{s2} are given by,

$$\begin{aligned} \mathcal{X}_{s1}(k) &= s_k, \\ \mathcal{X}_r(k) &= \begin{cases} s_k, & \text{if } k \text{ is even,} \\ s_{(k+M/2) \bmod M}, & \text{if } k \text{ is odd,} \end{cases} \\ \mathcal{X}_{s2}(k) &= s_k, \end{aligned}$$

where $0 \leq k \leq M-1$.

Theorem 2: For the labelling scheme $\tilde{\mathcal{L}}$, with 2^l -PSK signal set, the labelling gain, assuming $\gamma \ll 1$, is approximately given by,

$$L_G(\tilde{\mathcal{L}}) \approx \begin{cases} 20 \log_{10} [\cot(\pi/2^l)], & \text{for } l = 2, 3; \\ 20 \log_{10} [4 \cos^2(\pi/2^l)], & \text{for } l \geq 4. \end{cases} \quad (37)$$

Proof: Let us define,

$$\begin{aligned} d_1(k, k') &= |\mathcal{X}_{s1}(k) - \mathcal{X}_{s1}(k')| \\ &= |\exp(jk2\pi/M) - \exp(jk'2\pi/M)| \\ &= |1 - \exp(j(k' - k)2\pi/M)|, \end{aligned}$$

where $0 \leq k, k' \leq M-1$.

$$\begin{aligned}
Pr(a \rightarrow \bar{a}) \leq & \sum_{l=1}^M \left(\left[\frac{1}{1 + \frac{|\sigma_{ds}(x_{s_1}(a) - x_{s_1}(\bar{a}))|^2}{4}} \right] \left[\frac{1}{1 + \frac{|\sigma_{ds}(x_{s_2}(a) - x_{s_2}(\bar{a}))|^2}{4} + \frac{|\sigma_{dr}(x_r(a) - x_r(l))|^2}{4}} \right] \right. \\
& \left. \left[\frac{1}{1 + \frac{|\sigma_{rs}(x_{s_1}(\bar{a}) - x_{s_1}(l))|^2}{8}} \right] \right) \\
& + \sum_{j=1, j \neq a}^M \sum_{m=1}^M \left(\left[\frac{1}{1 + \frac{|\sigma_{ds}(x_{s_1}(a) - x_{s_1}(\bar{a}))|^2}{4}} \right] \left[\frac{1}{1 + \frac{|\sigma_{ds}(x_{s_2}(a) - x_{s_2}(\bar{a}))|^2}{4} + \frac{|\sigma_{dr}(x_r(j) - x_r(m))|^2}{4}} \right] \right. \\
& \left. \left[\frac{1}{1 + \frac{|\sigma_{rs}(x_{s_1}(a) - x_{s_1}(j))|^2}{8} + \frac{|\sigma_{rs}(x_{s_1}(\bar{a}) - x_{s_1}(m))|^2}{8}} \right] \right) \quad (36)
\end{aligned}$$

Since $d_1(k, k')$ depends only on $|k - k'|$, let us denote it by $d_1(n)$, where $0 \leq |k - k'| = n \leq M$. Then,

$$\begin{aligned}
d_1(n) &= |1 - \exp(jn2\pi/M)| \\
&= 2|\sin(\pi n/M)|, \quad (38)
\end{aligned}$$

where $0 \leq n \leq M$. Similarly, let us define,

$$\begin{aligned}
d_2(n) &= |\mathcal{X}_r(k) - \mathcal{X}_r(k')| \\
&= \begin{cases} |1 - \exp(jn2\pi/M)|, & \text{for } n \text{ even} \\ |1 + \exp(jn2\pi/M)|, & \text{for } n \text{ odd} \end{cases} \\
&= \begin{cases} 2|\sin(\pi n/M)|, & \text{for } n \text{ even} \\ 2|\cos(\pi n/M)|, & \text{for } n \text{ odd} \end{cases} \quad (39)
\end{aligned}$$

Since $\gamma \ll 1$, we have

$$\begin{aligned}
d(\bar{\mathcal{L}}) &\approx \min_{k \neq k'} |\mathcal{X}_{s_1}(k) - \mathcal{X}_{s_1}(k')|^2 |\mathcal{X}_r(k) - \mathcal{X}_r(k')|^2 \\
&= \min_{1 \leq n \leq M} d_1^2(n) d_2^2(n) \\
&= \min_{1 \leq n \leq M} (\min_{n \text{ even}} [d_1^2(n) d_2^2(n)], \min_{n \text{ odd}} [d_1^2(n) d_2^2(n)]). \quad (40)
\end{aligned}$$

From (38) and (39) we have,

$$\begin{aligned}
\min_{\substack{1 \leq n \leq M \\ n \text{ even}}} [d_1^2(n) d_2^2(n)] &= \min_{\substack{1 \leq n \leq M \\ n \text{ even}}} [16 \sin^4(\pi n/M)] \\
&= 16 \sin^4(2\pi/M) \quad (41)
\end{aligned}$$

and similarly,

$$\begin{aligned}
\min_{\substack{1 \leq n \leq M \\ n \text{ odd}}} [d_1^2(n) d_2^2(n)] &= \min_{\substack{1 \leq n \leq M \\ n \text{ odd}}} [16 \sin^2(\pi n/M) \cos^2(\pi n/M)] \\
&= 4 \sin^2(2\pi/M). \quad (42)
\end{aligned}$$

Substituting (41) and (42) in (40) we have,

$$d(\bar{\mathcal{L}}) \approx \min[16 \sin^4(2\pi/M), 4 \sin^2(2\pi/M)]. \quad (43)$$

Also, we have

$$\begin{aligned}
d(\mathcal{L}_0) &\approx \min_{1 \leq n \leq M} d_1^4(n) \\
&= 16 \sin^4(\pi/M). \quad (44)
\end{aligned}$$

Hence, from (43) and (44), by definition of $L_G(\bar{\mathcal{L}})$, we have,

$$L_G(\bar{\mathcal{L}}) = 20 \log_{10}[\min(\cot(\pi/M), 4 \cos^2(\pi/M))].$$

It can be verified that $\cot(\pi/M) \gtrsim 4 \cos^2(\pi/M)$ for $M \gtrsim 12$, from which we get (37). ■

It can be seen from Theorem 2 that for $l = 2$, the labelling gain $L_G(\bar{\mathcal{L}}) \approx 0$ dB. In other words, for the uncoded transmission scheme using 4 PSK, using different labelling schemes at S and R does not provide any performance improvement. It can be seen that $\cot(\pi/2^l)$ for $l = 3$, and $4 \cos^2(\pi/2^l)$ for $l \geq 4$ are greater than one, whence the approximate value of $L_G(\bar{\mathcal{L}})$ is greater than 0 dB for $l \geq 3$. Since $L_G(\bar{\mathcal{L}})$ is always greater than the approximate value given by Theorem 2, for the uncoded transmission scheme using 2^l PSK, using our labelling scheme provides advantage over constant labelling, for all $l \geq 3$. Furthermore, since $4 \cos^2(\pi/2^l)$ is an increasing function of l , the labelling gain increases with increasing size of the PSK constellation. Simulation results presented in Section VII confirm that the uncoded transmission with our labelling outperforms the uncoded transmission with constant labelling.

VI. TCM CODE DESIGN GUIDELINES AND EXAMPLES

A. TCM Code Design Guidelines

Let the generalized product distance $m_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ be split into two parts as,

$$m_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S) = m_{21}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) + m_{22}(\mathcal{P}_S, \tilde{\mathcal{P}}_S),$$

where the product distances $m_{21}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ and $m_{22}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ are given by,

$$m_{21}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) = \prod_{i \in \eta_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S)} \left| x_r^i(\mathcal{P}_S) - x_r^i(\tilde{\mathcal{P}}_S) \right|^2,$$

$$m_{22}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) = \prod_{i \in \eta_2(\mathcal{P}_S, \tilde{\mathcal{P}}_S)} \gamma \left| x_{s_2}^i(\mathcal{P}_S) - x_{s_2}^i(\tilde{\mathcal{P}}_S) \right|^2.$$

Let \mathcal{G}_1 and \mathcal{G}_2 be two metrics defined as,

$$\mathcal{G}_1 = \min_{(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \in \mathcal{Y}_P} \left[m_1(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \right],$$

$$\mathcal{G}_2 = \min_{(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \in \mathcal{Z}_P} \left[m_1(\mathcal{P}_S, \tilde{\mathcal{P}}_S) m_{21}(\mathcal{P}_S, \tilde{\mathcal{P}}_S) \right],$$

where \mathcal{Y}_P denotes the set of pairs of paths whose effective length is equal to the effective length of the trellis \mathcal{T}_{S_1} .

The code design guidelines are summarized below. Some of the design rules for TCM for the point to point AWGN and fading channels ([20], [21]), carry over for the relay channel as well.

- Since trellises with parallel transitions limit the effective length of the trellis to one, they should be avoided.
- Signal points from the signal set should occur with equal frequency.
- In a trellis with regularity and symmetry, two branches which emerge from the same state form a part of a pair of paths which differ by a unmerged length equal to the unmerged length of the code. Since the coding gain metric \mathcal{G} involves parameters $m_1(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, $m_{21}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$ and $m_{22}(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, which are of the form of product of the branch Euclidean distances of the pair of paths $(\mathcal{P}_S, \tilde{\mathcal{P}}_S)$, signal points assigned to branches emerging from the same state should be from the same Ungerboeck partition [20].

Since $\gamma \ll 1$, the values of \mathcal{G} and \mathcal{G}_2 are nearly equal and instead of maximizing \mathcal{G} , \mathcal{G}_2 can be maximized. In the design examples presented, the choice of \mathcal{X}_{s_1} is made such that the effective length of \mathcal{T}_{S_1} is maximized and the value of \mathcal{G}_1 is made large. The choice of \mathcal{X}_{s_2} and \mathcal{X}_r are made such that the effective length of \mathcal{T}_{S_2} is maximized and the value of \mathcal{G}_2 is made large. In the design examples presented, while the diversity order obtained is maximum, it is not claimed that the value of \mathcal{G}_2 for all the cases is maximum. As was the case with TCM for fading channel, maximizing the coding gain is a separate problem in itself, as it heavily depends on the trellis and the signal set used.

B. Code Design Examples

In all the trellis diagrams shown, the labellings on the edges are shown to the left of each state. The triple inside $[\cdot]$, when read from left to right, denotes the labellings $(\mathcal{X}_{s_1}, \mathcal{X}_r, \mathcal{X}_{s_2})$ corresponding to the edges emerging from top to bottom (the labellings $\mathcal{X}_{s_1}, \mathcal{X}_r, \mathcal{X}_{s_2}$ are shown in the same trellis diagram instead of three different trellis diagrams). In Examples 2, 3, 4 and 5 considered in this subsection, S transmits two information bits to D in the two phases of relaying.

Example 2: Consider the case where encoding at S and R take place using a two state trellis (Fig. 9) with 8 PSK signal

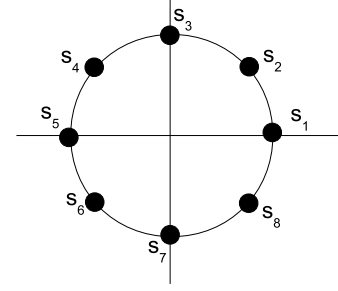


Fig. 8. 8-PSK signal set

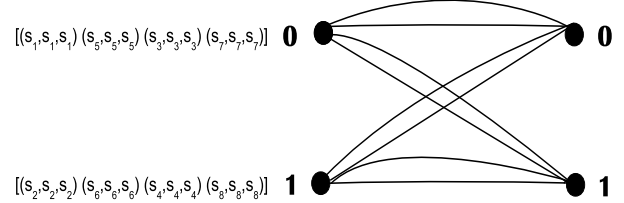


Fig. 9. Two State Trellis

set with signal points labelled as shown in Fig. 8. Let δ_0 , δ_1 , δ_2 and δ_3 denote the ordered squared Euclidean distances between the points in the 8 PSK signal set, i.e., $\delta_0 = (2 - \sqrt{2}) = 0.586$, $\delta_1 = 2$, $\delta_2 = (2 + \sqrt{2}) = 3.414$ and $\delta_3 = 4$. Since the unmerged length of the code is equal to one, the diversity order cannot exceed two. The choice of \mathcal{X}_{s_1} shown in Fig. 9 results in the value of the effective length of the trellis \mathcal{T}_{S_1} equal to one and maximizes \mathcal{G}_1 . The value of $\mathcal{G}_1 = \delta_3 = 4$. Choosing \mathcal{X}_r and \mathcal{X}_{s_2} same as \mathcal{X}_{s_1} results in the value of the effective length of \mathcal{T}_{S_2} equal to one and maximizes \mathcal{G}_2 . For this example, the diversity order is 2 and the coding gain metric $\mathcal{G} = \delta_3(\delta_3 + \gamma\delta_3) = 16.5056$, for $\gamma = 0.0316$.

Example 3: Consider the case where encoding at S and R take place using a four state trellis (Fig. 10) with 8 PSK signal set. The unmerged length of the code is two and hence the diversity order cannot exceed four. From [22], the choice of \mathcal{X}_{s_1} shown in Fig. 10 results in a value of the effective length of \mathcal{T}_{S_1} equal to two and $\mathcal{G}_1 = \delta_0\delta_3 = 2.344$.

If \mathcal{X}_{s_1} , \mathcal{X}_{s_2} and \mathcal{X}_r are chosen based on the design criteria for the fading channel (Jamali et al. labelling [22]), the value obtained for the metric $\mathcal{G}_2 = (\delta_0\delta_3)^2 = 5.49$. Instead, if the labellings \mathcal{X}_{s_1} , \mathcal{X}_{s_2} and \mathcal{X}_r are chosen as shown in Fig. 10 the value of $\mathcal{G}_2 = \delta_0\delta_1\delta_2\delta_3 = 16$ and the effective length of \mathcal{T}_{S_2} equal to two. For this example, the diversity order is 4 and the coding gain metric $\mathcal{G} = \delta_0\delta_3(\delta_1\delta_2 + \gamma\delta_0\delta_3) = 16.1735$, for $\gamma = 0.0316$.

Example 4: Consider the case where encoding at S and R take place using an eight state trellis (Fig. 11) with 8 PSK signal set. For this trellis, the diversity order cannot exceed

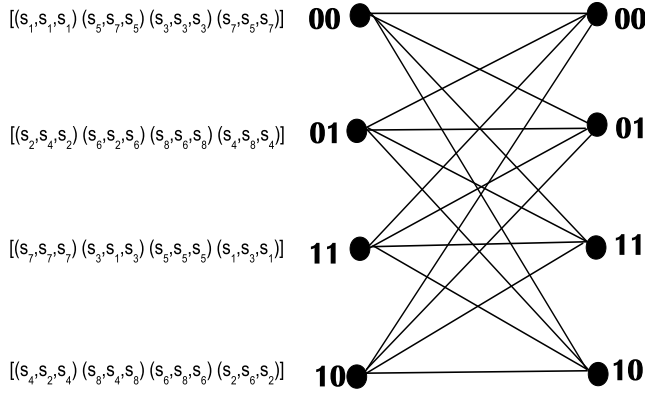


Fig. 10. Four State Trellis

four since the unmerged length of the code is two. From [23], the choice of \mathcal{X}_{s_1} shown in Fig. 11 (which is the same as Ungerboeck's labelling [20]) results in a value of the effective length equal to two and $\mathcal{G}_1 = \delta_1 \delta_3 = 8$. Choosing \mathcal{X}_r and \mathcal{X}_{s_2} same as \mathcal{X}_{s_1} results in $\mathcal{G}_2 = (\delta_1 \delta_3)^2 = 64$ and ensures that the effective length of in \mathcal{T}_{S_2} is equal to two. For this example, the diversity order is 4 and the coding gain metric $\mathcal{G} = (\delta_1 \delta_3)^2(1 + \gamma) = 66.0224$, for $\gamma = 0.0316$. Increasing the number of states from 4 to 8, while it provides an increase in the coding gain, does not provide diversity advantage.

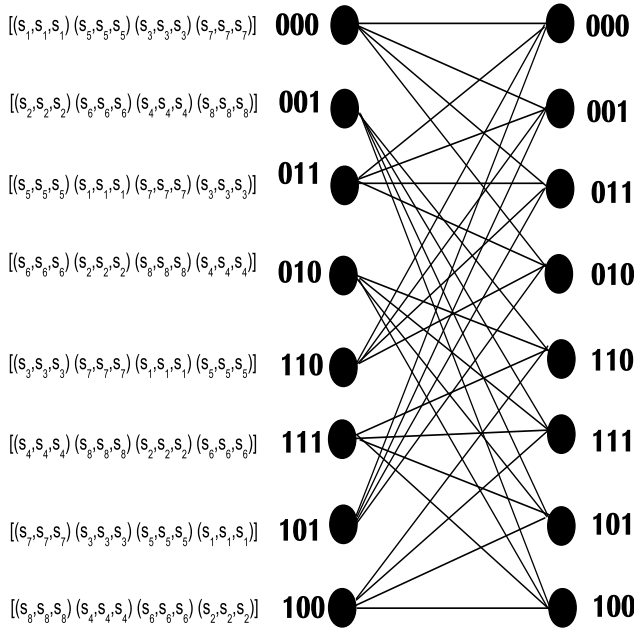


Fig. 11. 8 State Trellis

Example 5: Consider the case where encoding at S and R

take place using a sixteen state trellis shown in Fig. 12. The unmerged length of the code is three and hence the diversity order cannot exceed six. From [22], the choice of \mathcal{X}_{s_1} shown in Fig. 12 (which is the same as Ungerboeck's labelling [20]) results in a value of the effective length of \mathcal{T}_{S_1} equal to three and $\mathcal{G}_1 = \delta_0 \delta_1 \delta_3 = 4.68$. Labellings \mathcal{X}_r and \mathcal{X}_{s_2} chosen to be the same as \mathcal{X}_{s_1} result in $\mathcal{G}_2 = (\delta_0 \delta_1 \delta_3)^2 = 21.96$ and the effective length of \mathcal{T}_{S_2} equal to three. For this example, the diversity order is 6 and the coding gain metric $\mathcal{G} = (\delta_0 \delta_1 \delta_3)^2(1 + \gamma) = 22.66$, for $\gamma = 0.0316$.

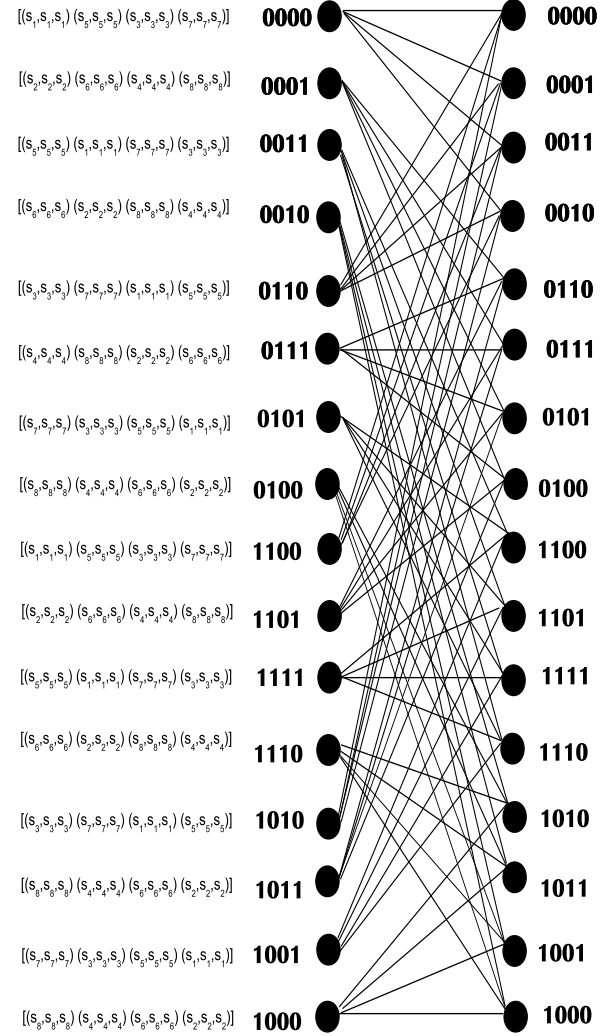


Fig. 12. 16 State Trellis

In the following example we consider the case where 3 bits of information get transmitted from S to D in the two phases of relaying.

Example 6: Consider the case where encoding at S and R take place using an eight state trellis (Fig. 14) with 16 PSK signal set whose points are labelled as shown in Fig. 13. For this trellis the diversity order cannot exceed four, since the unmerged length of the code is two. The choice of \mathcal{X}_{s_1} shown in Fig. 14 results in a value of $\mathcal{G}_1 = 0.0892$ and ensures that

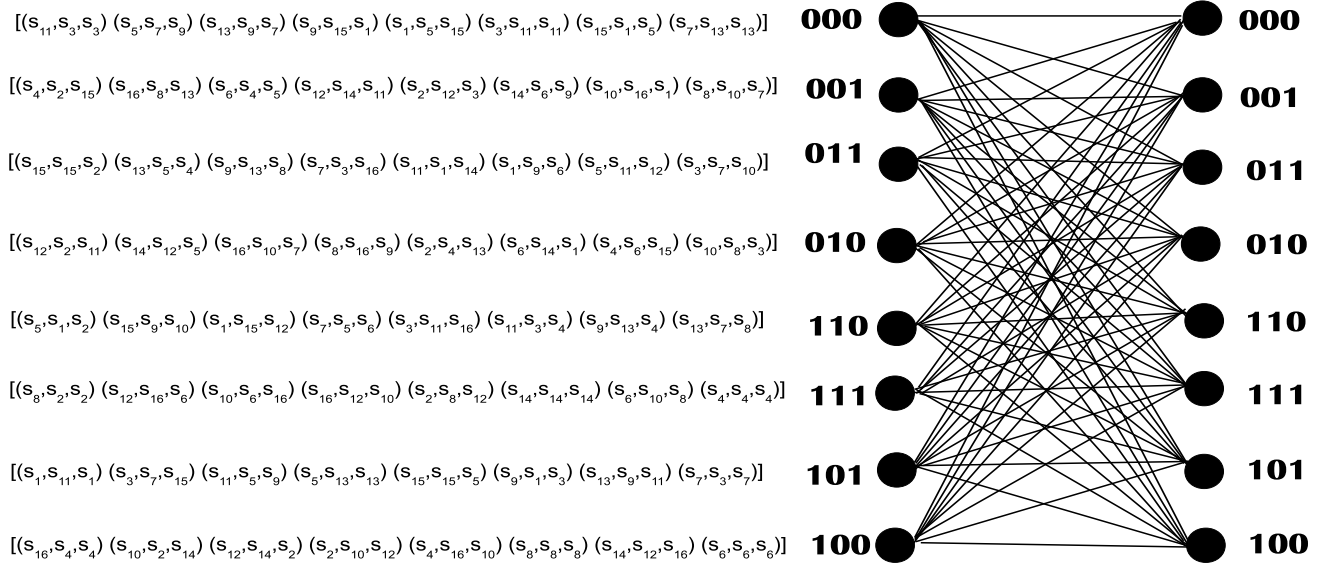


Fig. 14. 8 State Trellis

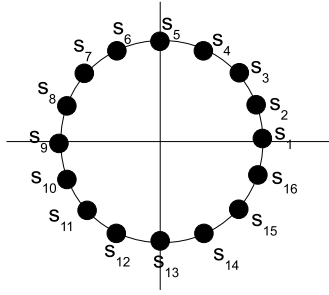


Fig. 13. 16 PSK Signal Set

the effective length of \mathcal{T}_{S_1} is two. The labellings \mathcal{X}_r and \mathcal{X}_{s_2} shown in Fig. 14 result in $\mathcal{G}_2 = 0.1177$ and ensures that the effective length of \mathcal{T}_{S_2} is two. For this example, the diversity order is 4 and the coding gain metric $\mathcal{G} = 0.1295$, for $\gamma = 0.1$.

VII. SIMULATION RESULTS

The E_S vs BER plots for the different schemes achieving a spectral efficiency of 1 bpcu are shown in Fig. 15, for $\sigma_{sd}^2 = 0$ dB, $\sigma_{sr}^2 = 15$ dB and $\sigma_{rd}^2 = 15$ dB. The diversity order and the value of E_S required at a BER of 10^{-4} are summarized in Table I. As observed in the previous section, the uncoded transmission scheme provides a diversity order 2. We observe that using a 2 state trellis provides a diversity order 2 as expected. At a BER of 10^{-4} , the 2 state 8 PSK TCM scheme provides 4.5 dB gain over the uncoded transmission scheme using 4 PSK. When the number of states is increased to 4,

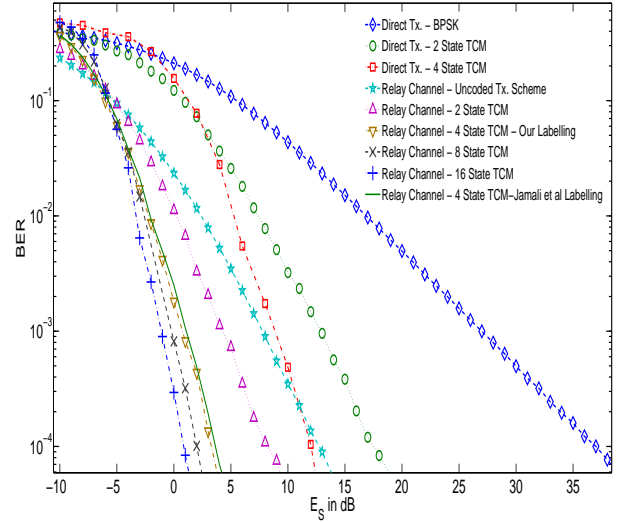


Fig. 15. E_S vs BER for the uncoded transmission scheme using 4 PSK and TCM Schemes for $\sigma_{sd}^2 = 0$ dB, $\sigma_{sr}^2 = 10$ dB and $\sigma_{rd}^2 = 10$ dB

we see that the diversity order increases to 4 as predicted and at a BER of 10^{-4} a large gain of 9.5 dB is obtained over the uncoded transmission scheme using 4 PSK. Also, it can be seen from Fig. 15, it can be seen that the 4 state 8 PSK scheme with our labelling provides a gain of 0.6 dB over the case when S and R use Jamali et al. labelling. Increasing the number of states to 8, from Fig. 15 the diversity order stays

TABLE II
COMPARISON OF UNCODED 8 PSK TRANSMISSION SCHEME AND 16 PSK TCM SCHEMES

Scheme	Diversity Order	E_S in dB for $BER = 10^{-4}$
Uncoded Transmission Scheme - 8 PSK - With Constant Labelling	2	20 dB
Uncoded Transmission Scheme - 8 PSK - With Our Labelling	2	18 dB
Relay Channel - 8 State TCM	4	12.5 dB

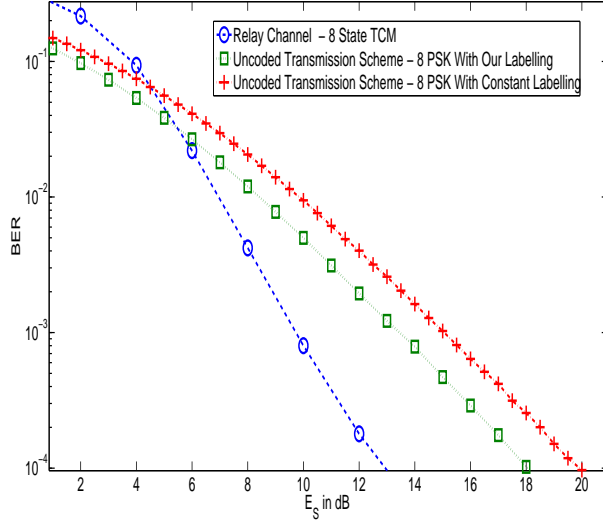


Fig. 16. E_S vs BER for the uncoded transmission scheme using 8 PSK and TCM Schemes for $\sigma_{sd}^2 = 0$ dB, $\sigma_{sr}^2 = 10$ dB and $\sigma_{rd}^2 = 10$ dB

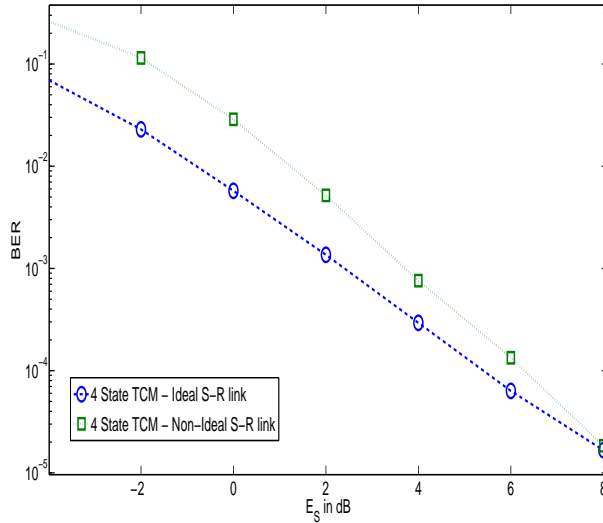


Fig. 17. E_S vs BER comparison for the proposed TCM scheme with ideal and non-ideal S-R links for $\sigma_{sd}^2 = 0$ dB, $\sigma_{sr}^2 = 10$ dB and $\sigma_{rd}^2 = 10$ dB

at 4 but a coding gain of 1.5 dB is obtained over the 4 state 8 PSK TCM scheme. With the 16 state 8 PSK TCM scheme, a diversity order of 6 is obtained and a gain of 1.5 dB is obtained over the 8 state 8 PSK TCM scheme. In essence, at a BER of 10^{-4} , a large gain of 12.5 dB is obtained using the 16 state 8 PSK TCM scheme over the uncoded transmission scheme using 4 PSK.

Fig. 16 shows the plots comparing the uncoded transmission scheme using 8 PSK with our labelling and with constant labelling, and the 8 state 16 PSK TCM scheme, for $\sigma_{sd}^2 = 0$ dB, $\sigma_{sr}^2 = 10$ dB and $\sigma_{rd}^2 = 10$ dB. At a BER of 10^{-4} the uncoded transmission scheme using 8 PSK with our labelling and 8 state 16 PSK TCM provide a gain of 2 dB and 7.5 dB respectively over the 8 PSK uncoded transmission scheme with constant labelling. The diversity order and the value of E_S required at a BER of 10^{-4} are summarized in Table II.

Fig. 17 shows the E_S vs BER plot for the 4 state 8 PSK TCM scheme, for the cases where the S-R link is ideal and non-ideal, for $\sigma_{sd}^2 = 0$ dB, $\sigma_{sr}^2 = 10$ dB and $\sigma_{rd}^2 = 10$ dB. As observed in Section IV E, the high SNR performance of the proposed near-ML decoder with non-ideal S-R link approaches the performance of the optimal ML decoder with an ideal S-R link.

VIII. DISCUSSION

A TCM scheme for the half duplex fading relay channel was proposed. A near-ML decoder whose high SNR performance approaches the performance of the optimal ML decoder was obtained. Based on the expression for PEP bounds, code design criteria to maximize the diversity order and coding gain were formulated. Maximizing the diversity order can be done by a proper choice of the labellings used at the source during Phase 1 and Phase 2. A unified procedure to find the labelling scheme which maximizes the coding gain for all signal sets and trellises is not known and will be an interesting topic for further research.

ACKNOWLEDGEMENT

This work was supported partly by the DRDO-IISc program on Advanced Research in Mathematical Engineering through a research grant as well as the INAE Chair Professorship grant to B. S. Rajan.

REFERENCES

- [1] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel", IEEE Trans. on Inform. Theory, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [2] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On the capacity of cheap relay networks", Conference on Information Sciences and Systems, Baltimore, USA, March 2003.
- [3] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On the capacity of cheap relay networks", IEEE GLOBECOM, USA, December 2003.

- [4] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels", *IEEE Trans. on Inform. Theory*, vol. 51, no. 6, pp. 2020-2040, June 2005.
- [5] Gerhard Kramer, Michael Gastpar and Piyush Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks", *IEEE Trans On Inform. Theory*, Vol. 51, NO. 9, Ssptember 2005.
- [6] Zhang Yong, Xu YouYun and Cai YueMing, "Power allocation for non-orthogonal decode and forward cooperation protocol", *Sci China Ser F-Inf Sci*, 2009, 52: 1037-1042.
- [7] R. Nabar, H. Bolcskei, and F. W. Kneubhler, "Fading relay channels: Performance limits and space-time signal design", *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1099-1109, Aug. 2004.
- [8] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity multiplexing tradeoff in half-duplex cooperative channels", *IEEE Trans. Inf. Theory*, vol. 51, pp. 4152-4172, Dec 2005.
- [9] P. Elia, K. Vinodh, M. Anand, and P. V. Kumar, "D-MG Tradeoff and Optimal Codes for a Class of AF and DF Cooperative Communication Protocols", *ISIT 2007*, pp. 681-685.
- [10] Z. Zhang and T. M. Duman, "Capacity-approaching turbo coding and iterative decoding for relay channels", *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1895-1905, Nov. 2005.
- [11] Z. Zhang and T. M. Duman, "Capacity-approaching turbo coding for half-duplex relaying", *IEEE Trans. Commun.*, vol. 55, no. 10, pp. 1895-1906, Oct. 2007.
- [12] Andrew Sendonaris, Elza Erkip and Behnaam Aazhang, "User Cooperation Diversity Part II: Implementation Aspects and Performance Analysis", *IEEE Transactions on Communications*, Vol. 51, No. 11, November 2003.
- [13] Tairan Wang, Alfonso Cano, Georgios B. Giannakis and J. Nicholas Laneman, "High-Performance Cooperative Demodulation With Decode-and-Forward Relays", *IEEE Transactions on Communications*, Vol. 55, No. 7, July 2007.
- [14] Deqiang Chen and J. Nicholas Laneman, "Modulation and Demodulation for Cooperative Diversity in Wireless Systems", *IEEE Transactions On Wireless Communications*, Vol. 5, No. 7, July 2006.
- [15] Xianglan Jin, Dong-Sup Jin, Jong-Seon No, Dong-Joon Shin, "On the Diversity Analysis of Decode-and-Forward Protocol With Multiple Antennas", *ISIT 2009*, Seoul, South Korea.
- [16] R. Tannious and A. Nosratinia, "Spectrally-efficient relay selection with limited feedback", *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1419-1428, Oct. 2008.
- [17] Xi Liu, Teng Joon Lim, "Fountain codes over fading relay channels", *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*, VOL. 8, NO. 6, JUNE 2009.
- [18] Jialing Li and Andrej Stefanov, "Cooperative Multiple Trellis Coded Modulation", *IEEE Transactions On Communications*, Vol. 57, No. 3, March 2009.
- [19] Vijayvaradharaj T. Muralidharan and B. Sundar Rajan, "Bounds on the Achievable Rate for the Fading Relay Channel with Finite Input Constellations," available online at arXiv: <http://arxiv.org/abs/1102.4272> [cs.IT], 21 Feb. 2011.
- [20] Ungerboeck G, "Channel coding with multilevel/phase signals", *IEEE Transactions On Information Theory*, Jan 1982.
- [21] Dariush Divsalar, "The Design of Trellis Coded MPSK for Fading Channels: Performance Criteria", *IEEE Transactions On Communications*, Vol. 36, No. 9, September 1988.
- [22] S. Hamidreza Jamali and Ngoc Tho Le, "Coded-modulation techniques for fading channels", Kluwer Academic Publishers 1994.
- [23] C. Schlegel and D. J. Costello, Jr., "Bandwidth efficient coding for fading channels: Code construction and performance analysis", *IEEE J. Select. Areas Commun.*, Vol. SAC-7, pp. 1356-1368, Dec. 1989.